STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 119466

To: Karen A Lacourciere **Location:** rem/2d15/2c18 **Art Unit:** 1635 **Tuesday, April 20, 2004** **Case Serial Number:** 09/310844	**From:** Beverly Shears **Location:** Remsen Bldg. RM 1A54 **Phone:** 571-272-2528 beverly.shears@uspto.gov

Search Notes

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 18, 2004, 07:30:54 ; Search time 1527.67 Seconds
(without alignments)
566.880 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 375216

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13	44.8	46	28	AZ833686	AZ833686 2M0115L20
	2	12.8	44.1	70	28	BH759592	BH759592 KG05236-3
c	3	12.8	44.1	72	13	BQ613481	BQ613481 rd07h06.y
c	4	12.8	44.1	76	28	AQ025263	AQ025263 EP(3)3081
c	5	12.4	42.8	52	9	AA700959	AA700959 zf87d10.s
c	6	12.4	42.8	70	9	AA468615	AA468615 ne08b04.s
	7	12.2	42.1	48	28	AZ503560	AZ503560 1M0343E21
	8	12.2	42.1	56	28	BZ770420	BZ770420 SALK_1433
	9	12.2	42.1	66	13	BX744082	BX744082 BX744082
c	10	12.2	42.1	67	14	CD946435	CD946435 REN_47 Ge
c	11	12.2	42.1	75	28	AZ453746	AZ453746 1M0255A23
c	12	11.8	40.7	40	9	AA975071	AA975071 on03d07.s
c	13	11.8	40.7	49	10	BE970036	BE970036 601680150
	14	11.8	40.7	51	29	CC516004	CC516004 CH240_361
c	15	11.8	40.7	63	29	CG563472	CG563472 OST186777
c	16	11.8	40.7	65	9	AA733449	AA733449 vt73h08.r
	17	11.8	40.7	70	9	AI767928	AI767928 wi99c01.x
c	18	11.8	40.7	70	28	BH216023	BH216023 1006039G0
	19	11.8	40.7	72	29	BX534001	BX534001 Arabidops
	20	11.8	40.7	74	29	CG574740	CG574740 OST207881
c	21	11.8	40.7	75	29	CNS01561	AL105043 Drosophil
	22	11.8	40.7	76	9	AV962932	AV962932 AV962932
c	23	11.8	40.7	76	28	BZ289518	BZ289518 SALK_0229
	24	11.8	40.7	77	13	BQ100875	BQ100875 ij25a03.x
	25	11.8	40.7	78	28	BH904967	BH904967 SALK_1054
c	26	11.8	40.7	79	29	CG573962	CG573962 OST206260
	27	11.8	40.7	79	29	AL951473	AL951473 Arabidops
	28	11.6	40.0	49	14	U44334	U44334 ENU44334 As
c	29	11.6	40.0	58	9	AI584456	AI584456 fb93h12.x
c	30	11.6	40.0	65	29	CG519587	CG519587 OST83436
	31	11.6	40.0	70	9	AI814489	AI814489 wj73g11.x
	32	11.6	40.0	73	29	BX001193	BX001193 Arabidops
	33	11.6	40.0	77	29	CG549254	CG549254 OST152329
c	34	11.6	40.0	77	29	AL759596	AL759596 Arabidops
c	35	11.6	40.0	79	14	CF115782	CF115782 SSFP015D1
c	36	11.6	40.0	79	28	AZ336769	AZ336769 1M0067L12
c	37	11.4	39.3	32	29	HSMC42B09	X88068 H.sapiens D
	38	11.4	39.3	34	29	BX001854	BX001854 Arabidops
c	39	11.4	39.3	40	29	TA253H01Q	AL483109 T. brucei
	40	11.4	39.3	46	9	AI887082	AI887082 wl96e09.x
	41	11.4	39.3	54	29	CC556861	CC556861 CH240_464
	42	11.4	39.3	61	14	CD946166	CD946166 REJ_33 Ge
c	43	11.4	39.3	61	14	CD963134	CD963134 SDS_4 Gen
c	44	11.4	39.3	66	9	AA247859	AA247859 j337l.seq
	45	11.4	39.3	67	29	TA113E04Q	AL460350 T. brucei
	46	11.4	39.3	68	9	AU254479	AU254479 AU254479
c	47	11.4	39.3	71	9	AU256177	AU256177 AU256177
	48	11.4	39.3	71	28	BH750951	BH750951 SALK_0465
c	49	11.4	39.3	73	9	AI900474	AI900474 sc11b10.y
c	50	11.4	39.3	75	28	BH609927	BH609927 HIV22C11
	51	11.4	39.3	75	29	CG526082	CG526082 OST102290
	52	11.4	39.3	76	9	AA988261	AA988261 os16a07.s
c	53	11.4	39.3	78	29	AG260396	AG260396 Lotus cor
	54	11.4	39.3	80	28	BZ357801	BZ357801 SALK_1312
	55	11.2	38.6	22	14	D18745	D18745 MUSGS01807
	56	11.2	38.6	30	9	AU259912	AU259912 AU259912
c	57	11.2	38.6	37	28	AZ950243	AZ950243 2M0214C15
	58	11.2	38.6	44	29	AL771575	AL771575 Arabidops
c	59	11.2	38.6	46	9	AA522160	AA522160 vf97c10.r
	60	11.2	38.6	49	9	AV841468	AV841468 AV841468
	61	11.2	38.6	54	28	BH224605	BH224605 1006120A0
	62	11.2	38.6	60	14	H19786	H19786 yn60a10.s1
	63	11.2	38.6	64	9	AI321110	AI321110 d4c09nm.r
c	64	11.2	38.6	65	29	AL755798	AL755798 Arabidops
	65	11.2	38.6	66	9	AI571487	AI571487 tr56g10.x
c	66	11.2	38.6	66	29	CG603111	CG603111 OST277306
c	67	11.2	38.6	67	9	AA936041	AA936041 nz53f10.s
	68	11.2	38.6	70	9	AU258214	AU258214 AU258214
c	69	11.2	38.6	70	14	CB916098	CB916098 VVD106F01
c	70	11.2	38.6	71	29	CG545852	CG545852 OST144556
	71	11.2	38.6	72	28	AZ452726	AZ452726 1M0252D13
	72	11.2	38.6	72	28	CC458528	CC458528 SALK_1194
c	73	11.2	38.6	73	9	AA865654	AA865654 og89b12.s
c	74	11.2	38.6	73	13	BX723784	BX723784 BX723784
c	75	11.2	38.6	75	28	AZ591678	AZ591678 1M0401I24
c	76	11.2	38.6	76	13	BQ240244	BQ240244 TaE05020B
	77	11.2	38.6	76	29	CG706108	CG706108 01S0707-0

c	78	11.2	38.6	77	14	CA845152 CA845152 hab98h11.
	79	11.2	38.6	77	14	CB262321 CB262321 66-E8867-
c	80	11.2	38.6	78	9	AA936218 AA936218 on43c10.s
c	81	11.2	38.6	79	28	AZ480000 AZ480000 1M0301H08
c	82	11.2	38.6	80	13	BX761812 BX761812 BX761812
c	83	11	37.9	34	29	TA98E04P AL461291 T. brucei
	84	11	37.9	40	9	AI016485 AI016485 ot86g10.s
	85	11	37.9	43	28	AZ639598 AZ639598 1M0501L07
c	86	11	37.9	43	29	BX659541 BX659541 Arabidops
c	87	11	37.9	56	12	BM445434 BM445434 1IL11C9.a
	88	11	37.9	75	10	AW516157 AW516157 xt61g04.x
c	89	11	37.9	75	28	BH617610 BH617610 SALK_0373
	90	11	37.9	75	29	AL754690 AL754690 Arabidops
	91	11	37.9	75	29	AL754692 AL754692 Arabidops
	92	11	37.9	77	28	BH252676 BH252676 SALK_0137
c	93	10.8	37.2	31	9	AA206191 AA206191 zq54a06.r
c	94	10.8	37.2	33	9	AU256066 AU256066 AU256066
	95	10.8	37.2	40	29	CG779591 1123034H1
c	96	10.8	37.2	44	28	BH811570 BH811570 SALK_0591
	97	10.8	37.2	45	29	BX289611 BX289611 Arabidops
	98	10.8	37.2	49	9	AI273069 AI273069 qv62g02.x
c	99	10.8	37.2	50	28	AZ817325 AZ817325 2M0086C23
	100	10.8	37.2	51	28	AZ983852 AZ983852 2M0265H09
c	101	10.8	37.2	51	28	BH232806 1006169G0
	102	10.8	37.2	51	29	BX894095 BX894095 Arabidops
c	103	10.8	37.2	53	14	CB409428 CB409428 NISC_nc05
	104	10.8	37.2	53	29	AL940874 AL940874 Arabidops
c	105	10.8	37.2	56	9	AA610958 AA610958 AY386-2 A
	106	10.8	37.2	56	28	BZ354770 BZ354770 SALK_1257
c	107	10.8	37.2	56	28	BZ380049 BZ380049 SALK_1145
	108	10.8	37.2	58	29	AL948765 AL948765 Arabidops
	109	10.8	37.2	59	14	CB258816 CB258816 12-E01409
c	110	10.8	37.2	61	13	BQ479345 BQ479345 ku33d12.y
c	111	10.8	37.2	61	14	T12612 T12612 CHR90132 Ch
	112	10.8	37.2	61	28	B05670 B05670 cSRL-69f2-u
	113	10.8	37.2	61	28	BH901722 BH901722 SALK_0858
c	114	10.8	37.2	61	29	CG523070 CG523070 OST93955
c	115	10.8	37.2	62	9	AU009292 AU009292 AU009292
	116	10.8	37.2	62	29	CG522033 CG522033 OST90954
	117	10.8	37.2	63	9	AA668218 AA668218 ab77d07.s
c	118	10.8	37.2	63	12	BG362434 BG362434 gb72b09.y
c	119	10.8	37.2	64	9	AA179842 AA179842 zp53f06.s
c	120	10.8	37.2	64	29	BX291457 BX291457 Arabidops
c	121	10.8	37.2	65	28	AZ329128 AZ329128 1M0053B17
c	122	10.8	37.2	65	29	CG502754 CG502754 OST48383
c	123	10.8	37.2	66	14	CD925111 CD925111 G750.115O
	124	10.8	37.2	67	29	AL942571 AL942571 Arabidops
c	125	10.8	37.2	69	10	BE647308 BE647308 UI-M-BH1-
c	126	10.8	37.2	69	29	CG617320 CG617320 OST310685
	127	10.8	37.2	70	28	AZ049151 AZ049151 GSSBru055
c	128	10.8	37.2	70	28	AZ778222 AZ778222 2M0013H19
	129	10.8	37.2	70	29	CG529233 CG529233 OST109469
c	130	10.8	37.2	71	14	CK109829 CK109829 N029E03 P
c	131	10.8	37.2	71	29	CG638699 CG638699 OST368709
	132	10.8	37.2	71	29	CG646383 CG646383 OST392555
c	133	10.8	37.2	71	29	AL949476 AL949476 Arabidops
c	134	10.8	37.2	72	9	AA570253 AA570253 nf39e03.s
	135	10.8	37.2	72	10	AW116071 AW116071 fi06g06.x
c	136	10.8	37.2	73	28	AZ943484 AZ943484 2M0204K09
c	137	10.8	37.2	75	12	BI944911 BI944911 sag27b05.
	138	10.8	37.2	75	14	CK098783 CK098783 A031P73.5
	139	10.8	37.2	76	28	BZ291884 BZ291884 SALK_1219
c	140	10.8	37.2	76	29	CG630076 CG630076 OST344981
c	141	10.8	37.2	77	9	AA513220 AA513220 nh78g01.s
c	142	10.8	37.2	77	28	BH907776 BH907776 SALK_0440
	143	10.8	37.2	77	29	CC886985 CC886985 SALK_1493
c	144	10.8	37.2	78	29	BX650260 BX650260 Arabidops
	145	10.8	37.2	80	9	AU259251 AU259251 AU259251
	146	10.8	37.2	80	29	CG631171 CG631171 OST347263
c	147	10.6	36.6	29	28	AZ871142 AZ871142 2M0183E21
c	148	10.6	36.6	34	28	AZ840876 AZ840876 2M0138C08
	149	10.6	36.6	37	29	AL951243 AL951243 Arabidops
	150	10.6	36.6	39	14	H55495 H55495 CHR220434 C
	151	10.6	36.6	40	28	BZ595436 BZ595436 SALK_0870
	152	10.6	36.6	42	28	AZ760157 AZ760157 1M0553G12
	153	10.6	36.6	43	28	AZ484548 AZ484548 1M0311N02
c	154	10.6	36.6	43	28	AZ597048 AZ597048 1M0410K10
c	155	10.6	36.6	45	28	AZ502122 AZ502122 1M0341F14
	156	10.6	36.6	45	28	BZ595896 BZ595896 SALK_0897
c	157	10.6	36.6	46	9	AI900599 AI900599 sc13e01.y
	158	10.6	36.6	49	28	BH856818 BH856818 SALK_0791
	159	10.6	36.6	50	9	AU105844 AU105844 AU105844
	160	10.6	36.6	50	28	AZ304992 AZ304992 1M0005D12
c	161	10.6	36.6	52	29	AL761298 AL761298 Arabidops
c	162	10.6	36.6	53	28	AZ780230 AZ780230 2M0017H19
	163	10.6	36.6	56	28	AZ783727 AZ783727 2M0025N09
	164	10.6	36.6	57	14	CD381570 CD381570 PTMM06739
c	165	10.6	36.6	58	28	AZ834846 AZ834846 2M0117F18
	166	10.6	36.6	59	28	BZ762290 BZ762290 SALK_0978
	167	10.6	36.6	60	10	BF634030 BF634030 NF072H10D
c	168	10.6	36.6	61	9	AW107310 AW107310 um14c01.x
	169	10.6	36.6	61	14	CD966715 CD966715 SEQ_117 G
	170	10.6	36.6	62	13	C21586 C21586 HUMGS001059
	171	10.6	36.6	62	29	BX656049 BX656049 Arabidops
c	172	10.6	36.6	63	9	AJ443201 AJ443201 AJ443201
c	173	10.6	36.6	64	9	AI181092 AI181092 uh91g06.r
	174	10.6	36.6	64	9	AA388420 AA388420 vc95a03.r
c	175	10.6	36.6	64	10	BE588193 BE588193 SWOvAMCAQ
c	176	10.6	36.6	64	10	BE636252 BE636252 SWOvAMCAQ
c	177	10.6	36.6	65	9	AI019809 AI019809 ua93c08.r
c	178	10.6	36.6	65	29	AL763793 AL763793 Arabidops
c	179	10.6	36.6	66	12	BM397621 BM397621 5009-0-35
c	180	10.6	36.6	66	14	CF044340 CF044340 QCJ28f03.
c	181	10.6	36.6	66	28	AZ492869 AZ492869 1M0327B13
c	182	10.6	36.6	68	28	AZ575904 AZ575904 AST-T22B0
	183	10.6	36.6	68	28	AZ624741 AZ624741 1M0463D14
	184	10.6	36.6	68	29	CG627325 CG627325 OST337112
c	185	10.6	36.6	69	9	AI211081 AI211081 n0e06a1.f
	186	10.6	36.6	69	14	CF316149 CF316149 HD--05-F0
	187	10.6	36.6	70	9	AA940574 AA940574 vz46f08.r
c	188	10.6	36.6	70	29	CG551217 CG551217 OST158035
	189	10.6	36.6	71	29	AG264184 AG264184 Lotus cor
c	190	10.6	36.6	72	13	BX679171 BX679171 BX679171
c	191	10.6	36.6	72	28	AZ404115 AZ404115 1M0172B16
	192	10.6	36.6	72	28	AZ810735 AZ810735 2M0076B04
c	193	10.6	36.6	72	29	CG519016 CG519016 OST81596
	194	10.6	36.6	72	29	AL764829 AL764829 Arabidops
	195	10.6	36.6	73	12	BM568281 BM568281 sal02d02.
c	196	10.6	36.6	75	9	AI221576 AI221576 qg15f05.x
c	197	10.6	36.6	75	14	U44278 U44278 ENU44278 As
c	198	10.6	36.6	75	28	BH910959 BH910959 SALK_0636
c	199	10.6	36.6	76	9	AA966778 AA966778 s9b04a1.f
c	200	10.6	36.6	76	9	AI590908 AI590908 tw26g12.x
c	201	10.6	36.6	76	9	AU178735 AU178735 AU178735
	202	10.6	36.6	76	10	AW396272 AW396272 sh26e01.y
c	203	10.6	36.6	76	10	BE876362 BE876362 601486683
c	204	10.6	36.6	76	28	AZ854788 AZ854788 2M0158K05
c	205	10.6	36.6	76	29	CG648093 CG648093 OST400044
	206	10.6	36.6	77	9	AL934676 AL934676 AL934676
	207	10.6	36.6	77	13	BQ907390 BQ907390 P005D07 O
	208	10.6	36.6	77	14	CF013238 CF013238 QBK3e08.x
	209	10.6	36.6	77	28	BH907776 BH907776 SALK_0440
	210	10.6	36.6	78	9	AI202456 AI202456 qs67d01.x
c	211	10.6	36.6	78	12	BM015178 BM015178 603641172
	212	10.6	36.6	78	14	CK108462 CK108462 I018P45 P
c	213	10.6	36.6	78	29	AL938849 AL938849 Arabidops
	214	10.6	36.6	78	29	BX658642 BX658642 Arabidops
c	215	10.6	36.6	79	9	AI006020 AI006020 ua85a03.r
c	216	10.6	36.6	79	9	AI271934 AI271934 qj88f04.x
	217	10.6	36.6	79	10	BG045244 BG045244 saa39d12.
	218	10.6	36.6	79	14	W89294 W89294 mf62h03.r1
c	219	10.6	36.6	79	29	CG661339 CG661339 OST442356
c	220	10.6	36.6	80	14	CF043983 CF043983 QCJ23g09.
c	221	10.6	36.6	80	28	BH219090 BH219090 1006084D0
	222	10.6	36.6	80	29	AL770832 AL770832 Arabidops
c	223	10.6	36.6	80	29	AL943121 AL943121 Arabidops

c 224	10.4	35.9	25	28	AZ993079	AZ993079 2M0277P20
c 225	10.4	35.9	27	29	CG723079	CG723079 1119074F0
226	10.4	35.9	30	28	BZ385302	BZ385302 SALK_1370
c 227	10.4	35.9	37	28	BZ378465	BZ378465 SALK_1081
c 228	10.4	35.9	37	28	BZ665596	BZ665596 KG06771-3
c 229	10.4	35.9	40	9	AI583940	AI583940 ts08e06.x
230	10.4	35.9	46	28	BH792322	BH792322 SALK_0634
231	10.4	35.9	49	29	BX287070	BX287070 Arabidops
232	10.4	35.9	50	28	AZ817068	AZ817068 2M0086C07
233	10.4	35.9	51	29	AL755958	AL755958 Arabidops
c 234	10.4	35.9	52	9	AA142353	AA142353 ms08a01.r
c 235	10.4	35.9	52	29	AG233473	AG233473 Lotus cor
c 236	10.4	35.9	53	9	AA142402	AA142402 ms08g01.r
237	10.4	35.9	54	28	AZ637547	AZ637547 1M0497C02
c 238	10.4	35.9	54	29	CC882265	CC882265 01S0555-0
c 239	10.4	35.9	55	29	BX132064	BX132064 Danio rer
240	10.4	35.9	58	29	CG706225	CG706225 02S2019-0
241	10.4	35.9	59	9	AF052489	AF052489 AF052489
242	10.4	35.9	60	14	CB225131	CB225131 1OM29C12
c 243	10.4	35.9	60	29	AL767399	AL767399 Arabidops
244	10.4	35.9	61	28	BZ380982	BZ380982 SALK_1160
c 245	10.4	35.9	62	12	BI518174	BI518174 603041972
c 246	10.4	35.9	63	29	CG894375	CG894375 03S3061-0
c 247	10.4	35.9	64	9	AI247119	AI247119 qx52f10.x
c 248	10.4	35.9	64	10	BF228778	BF228778 SWOvL3CAN
c 249	10.4	35.9	64	29	CG709612	CG709612 1119014A0
250	10.4	35.9	65	28	BH414068	BH414068 1007036B0
251	10.4	35.9	66	28	BH631033	BH631033 1007096C0
c 252	10.4	35.9	66	29	CG773643	CG773643 1123013C1
253	10.4	35.9	67	9	AI708576	AI708576 as98f04.x
c 254	10.4	35.9	67	29	CG895668	CG895668 03S4740-0
c 255	10.4	35.9	68	9	AA872087	AA872087 oi12c05.s
256	10.4	35.9	69	14	CA938762	CA938762 sav37e08.
257	10.4	35.9	69	14	CD940509	CD940509 RAN_66 Ge
258	10.4	35.9	69	14	R05841	R05841 ye88e09.r1
259	10.4	35.9	70	9	AI833026	AI833026 at74d05.x
c 260	10.4	35.9	70	28	AZ693627	AZ693627 AST-1HBG2
c 261	10.4	35.9	72	9	AU008441	AU008441 AU008441
c 262	10.4	35.9	72	29	CG649015	CG649015 OST403570
c 263	10.4	35.9	73	29	BX572505	BX572505 Arabidops
c 264	10.4	35.9	73	29	BX572506	BX572506 Arabidops
265	10.4	35.9	74	28	AZ986310	AZ986310 2M0268O24
c 266	10.4	35.9	74	29	CG918140	CG918140 CH240_139
c 267	10.4	35.9	74	29	TA251B07Q	AL482221 T. brucei
268	10.4	35.9	76	9	AI300666	AI300666 qo22a12.x
c 269	10.4	35.9	76	10	BE867849	BE867849 601443622
270	10.4	35.9	76	29	AL757078	AL757078 Arabidops
271	10.4	35.9	77	29	CG617517	CG617517 OST311200
272	10.4	35.9	78	9	AL789031	AL789031 AL789031
c 273	10.4	35.9	79	14	CK085025	CK085025 camcoP001
c 274	10.4	35.9	79	29	CG466861	CG466861 01S0714-0
275	10.4	35.9	79	29	CG729461	CG729461 1119112E0
276	10.4	35.9	80	13	BU649505	BU649505 1112080G1
c 277	10.2	35.2	27	28	BH909838	BH909838 SALK_0561
c 278	10.2	35.2	29	28	BH866445	BH866445 SALK_1013
279	10.2	35.2	32	28	AZ469379	AZ469379 1M0282P14
c 280	10.2	35.2	32	28	AZ766102	AZ766102 1M0563I24
281	10.2	35.2	35	29	AL939013	AL939013 Arabidops
282	10.2	35.2	36	28	AZ469569	AZ469569 1M0283F07
c 283	10.2	35.2	37	9	AI521252	AI521252 to66h08.x
c 284	10.2	35.2	37	28	AZ992335	AZ992335 2M0276O23
285	10.2	35.2	38	29	AL767806	AL767806 Arabidops
c 286	10.2	35.2	42	28	BZ585503	BZ585503 3590_1_35
287	10.2	35.2	42	29	BX547447	BX547447 Arabidops
288	10.2	35.2	43	28	BH864639	BH864639 SALK_0965
c 289	10.2	35.2	44	28	BH850508	BH850508 SALK_0713
290	10.2	35.2	46	9	AA120160	AA120160 mn33c12.r
c 291	10.2	35.2	46	28	AZ307757	AZ307757 1M0010E04
292	10.2	35.2	46	28	AZ815494	AZ815494 2M0083G18
c 293	10.2	35.2	46	28	BZ761728	BZ761728 SALK_0777
c 294	10.2	35.2	47	28	AZ514477	AZ514477 1M036ID21
c 295	10.2	35.2	48	9	AU257445	AU257445 AU257445
296	10.2	35.2	49	9	AA661516	AA661516 nr18e12.s
297	10.2	35.2	50	9	AU104386	AU104386 AU104386
298	10.2	35.2	50	9	AU105319	AU105319 AU105319
299	10.2	35.2	50	9	AU105321	AU105321 AU105321
300	10.2	35.2	50	9	AU105323	AU105323 AU105323
301	10.2	35.2	50	9	AU105324	AU105324 AU105324
302	10.2	35.2	50	9	AU105325	AU105325 AU105325
303	10.2	35.2	50	9	AU105326	AU105326 AU105326
304	10.2	35.2	50	9	AU105327	AU105327 AU105327
305	10.2	35.2	50	9	AU105329	AU105329 AU105329
306	10.2	35.2	50	9	AU105330	AU105330 AU105330
307	10.2	35.2	50	9	AU105331	AU105331 AU105331
308	10.2	35.2	50	9	AU105332	AU105332 AU105332
309	10.2	35.2	50	9	AU105333	AU105333 AU105333
310	10.2	35.2	50	9	AU105334	AU105334 AU105334
311	10.2	35.2	50	9	AU105335	AU105335 AU105335
312	10.2	35.2	50	9	AU105337	AU105337 AU105337
313	10.2	35.2	50	9	AU105338	AU105338 AU105338
314	10.2	35.2	50	9	AU105339	AU105339 AU105339
315	10.2	35.2	50	9	AU105340	AU105340 AU105340
316	10.2	35.2	50	9	AU105341	AU105341 AU105341
317	10.2	35.2	50	9	AU105342	AU105342 AU105342
318	10.2	35.2	50	9	AU105343	AU105343 AU105343
319	10.2	35.2	50	9	AU105344	AU105344 AU105344
320	10.2	35.2	50	9	AU105345	AU105345 AU105345
321	10.2	35.2	50	9	AU105346	AU105346 AU105346
322	10.2	35.2	50	9	AU105347	AU105347 AU105347
323	10.2	35.2	50	9	AU105348	AU105348 AU105348
324	10.2	35.2	50	9	AU105350	AU105350 AU105350
325	10.2	35.2	50	9	AU105351	AU105351 AU105351
326	10.2	35.2	50	9	AU105352	AU105352 AU105352
327	10.2	35.2	50	9	AU105353	AU105353 AU105353
328	10.2	35.2	50	9	AU105355	AU105355 AU105355
329	10.2	35.2	50	9	AU105356	AU105356 AU105356
330	10.2	35.2	50	9	AU105363	AU105363 AU105363
331	10.2	35.2	50	9	AU105364	AU105364 AU105364
332	10.2	35.2	50	9	AU105365	AU105365 AU105365
333	10.2	35.2	50	9	AU105366	AU105366 AU105366
334	10.2	35.2	50	9	AU105367	AU105367 AU105367
335	10.2	35.2	50	9	AU105369	AU105369 AU105369
336	10.2	35.2	50	9	AU105370	AU105370 AU105370
337	10.2	35.2	50	9	AU105373	AU105373 AU105373
338	10.2	35.2	50	9	AU105374	AU105374 AU105374
339	10.2	35.2	50	9	AU105375	AU105375 AU105375
340	10.2	35.2	50	9	AU105376	AU105376 AU105376
341	10.2	35.2	50	9	AU105377	AU105377 AU105377
342	10.2	35.2	50	9	AU105378	AU105378 AU105378
343	10.2	35.2	50	9	AU105379	AU105379 AU105379
344	10.2	35.2	50	9	AU105380	AU105380 AU105380
345	10.2	35.2	50	9	AU105381	AU105381 AU105381
346	10.2	35.2	50	9	AU105382	AU105382 AU105382
347	10.2	35.2	50	9	AU105383	AU105383 AU105383
348	10.2	35.2	50	9	AU105384	AU105384 AU105384
349	10.2	35.2	50	9	AU105385	AU105385 AU105385
350	10.2	35.2	50	9	AU105386	AU105386 AU105386
351	10.2	35.2	50	9	AU105387	AU105387 AU105387
352	10.2	35.2	50	9	AU105388	AU105388 AU105388
353	10.2	35.2	50	9	AU105391	AU105391 AU105391
354	10.2	35.2	50	9	AU105392	AU105392 AU105392
355	10.2	35.2	50	9	AU105394	AU105394 AU105394
356	10.2	35.2	50	9	AU105465	AU105465 AU105465
357	10.2	35.2	50	9	AU106680	AU106680 AU106680
358	10.2	35.2	50	28	AZ800436	AZ800436 2M0058L20
c 359	10.2	35.2	51	12	BG361927	BG361927 gb49d10.y
c 360	10.2	35.2	51	14	R82121	R82121 18G8 Chromo
c 361	10.2	35.2	51	28	BH904072	BH904072 SALK_1039
362	10.2	35.2	51	29	CC887057	CC887057 SALK_1494
c 363	10.2	35.2	51	29	AL758037	AL758037 Arabidops
c 364	10.2	35.2	52	9	AA165996	AA165996 ms23b03.r
c 365	10.2	35.2	52	9	AA490834	AA490834 aa49h04.s
366	10.2	35.2	52	28	B02727	B02727 cSRL-159e9-
c 367	10.2	35.2	53	12	BG315130	BG315130 PO1.0.79
c 368	10.2	35.2	53	28	BH855159	BH855159 SALK_0865
c 369	10.2	35.2	55	9	AA903873	AA903873 oe78b02.s

c 370	10.2	35.2	55	9	AI439306	AI439306 ti59e04.x
371	10.2	35.2	55	12	BG093858	BG093858 uu91b01.x
c 372	10.2	35.2	55	28	BZ382010	BZ382010 SALK_1177
373	10.2	35.2	56	14	CF875659	CF875659 tric039xa
374	10.2	35.2	56	14	CK151344	CK151344 GS1-014 S
c 375	10.2	35.2	56	29	CG799256	CG799256 1118001E0
376	10.2	35.2	57	13	BQ548436	BQ548436 rd10e10.y
c 377	10.2	35.2	59	12	BG409161	BG409161 gb87b05.y
378	10.2	35.2	59	14	CF973364	CF973364 PSU_bithr
379	10.2	35.2	59	29	AL766549	AL766549 Arabidops
c 380	10.2	35.2	60	9	AI719784	AI719784 as41h08.x
381	10.2	35.2	60	9	AL588055	AL588055 AL588055
382	10.2	35.2	60	28	BH863764	BH863764 SALK_0945
383	10.2	35.2	61	9	AU007077	AU007077 AU007077
c 384	10.2	35.2	61	14	R70336	R70336 yj81d09.r1
c 385	10.2	35.2	61	29	CG590967	CG590967 OST245023
c 386	10.2	35.2	62	12	BG611715	BG611715 602613078
387	10.2	35.2	62	14	CD939981	CD939981 RAA_65 Ge
c 388	10.2	35.2	62	28	AZ767834	AZ767834 1M0567O23
c 389	10.2	35.2	62	29	CG475777	CG475777 OST4358 M
390	10.2	35.2	63	28	AZ920429	AZ920429 1006019G1
c 391	10.2	35.2	63	28	BH415699	BH415699 1007044D0
c 392	10.2	35.2	63	29	CG475498	CG475498 OST3832 M
393	10.2	35.2	64	9	AI570111	AI570111 tr74d11.x
c 394	10.2	35.2	64	10	AW874915	AW874915 SWYACAL04
395	10.2	35.2	64	10	BE588185	BE588185 SWOvAMCAQ
c 396	10.2	35.2	64	14	H50434	H50434 yo29e11.s1
397	10.2	35.2	64	28	AZ438245	AZ438245 1M0228A17
398	10.2	35.2	64	28	AZ920334	AZ920334 1006019D0
c 399	10.2	35.2	65	14	D25832	D25832 HUMGS04202
c 400	10.2	35.2	65	14	H39336	H39336 DR14 IFNgam
401	10.2	35.2	65	28	AZ826744	AZ826744 2M0102O11
c 402	10.2	35.2	65	28	BH855666	BH855666 SALK_0848
403	10.2	35.2	65	29	CG546761	CG546761 OST146689
c 404	10.2	35.2	66	9	AI311426	AI311426 ta45b05.x
c 405	10.2	35.2	66	12	BI786902	BI786902 sai54g08.
c 406	10.2	35.2	66	14	CF022538	CF022538 QBQ16f01.
407	10.2	35.2	66	28	BZ358009	BZ358009 SALK_1317
c 408	10.2	35.2	67	9	AA708911	AA708911 zl64a10.s
409	10.2	35.2	67	9	AA756835	AA756835 vu20a08.r
410	10.2	35.2	67	9	AJ396243	AJ396243 AJ396243
c 411	10.2	35.2	67	12	BG362558	BG362558 gb74c12.y
412	10.2	35.2	67	14	CF973566	CF973566 PSU_bltwo
c 413	10.2	35.2	67	14	H25196	H25196 ym56e04.r1
414	10.2	35.2	67	28	BH855810	BH855810 SALK_0845
c 415	10.2	35.2	67	29	CG624461	CG624461 OST328242
c 416	10.2	35.2	67	29	CG670839	CG670839 OST471710
417	10.2	35.2	68	9	AA551800	AA551800 nk04a12.s
c 418	10.2	35.2	68	12	BG523223	BG523223 3-27 Stev
c 419	10.2	35.2	68	28	AZ367374	AZ367374 1M0117E05
c 420	10.2	35.2	69	9	AU011957	AU011957 AU011957
421	10.2	35.2	69	13	BQ519774	BQ519774 rd01a05.y
422	10.2	35.2	69	28	BH256492	BH256492 KG03686-5
c 423	10.2	35.2	69	29	CG552987	CG552987 OST165018
c 424	10.2	35.2	69	29	AL944091	AL944091 Arabidops
425	10.2	35.2	70	9	AI305284	AI305284 qm08g12.x
c 426	10.2	35.2	70	9	AI609394	AI609394 tw93b03.x
427	10.2	35.2	70	9	AA429013	AA429013 zw19f03.s
428	10.2	35.2	70	14	N56229	N56229 J9323F Huma
429	10.2	35.2	70	14	W20330	W20330 zb44d01.r1
430	10.2	35.2	70	28	BZ354128	BZ354128 SALK_1232
431	10.2	35.2	70	29	CC795791	CC795791 SALK_0883
c 432	10.2	35.2	70	29	CG572129	CG572129 OST202362
c 433	10.2	35.2	71	28	BH228382	BH228382 1006146C1
c 434	10.2	35.2	72	14	CB915847	CB915847 VVD103F09
435	10.2	35.2	72	14	T25592	T25592 EST00628 Un
c 436	10.2	35.2	73	12	BG361878	BG361878 gb46b10.y
c 437	10.2	35.2	73	12	BI107307	BI107307 602894328
438	10.2	35.2	73	14	CB832621	CB832621 SWWbmfCAV
439	10.2	35.2	73	28	AZ489337	AZ489337 1M0321P09
440	10.2	35.2	73	28	AZ920626	AZ920626 1006020G0
c 441	10.2	35.2	74	13	BQ275575	BQ275575 pj36f10.y
c 442	10.2	35.2	75	9	AU076453	AU076453 AU076453
443	10.2	35.2	75	12	BI201837	BI201837 r4e03fs.r
c 444	10.2	35.2	75	29	CG480059	CG480059 OST11622
c 445	10.2	35.2	75	29	CG669655	CG669655 OST466791
c 446	10.2	35.2	76	9	AI583160	AI583160 tr98d01.x
447	10.2	35.2	76	14	CF332220	CF332220 NACL--08-
c 448	10.2	35.2	76	28	BH848469	BH848469 SALK_0683
449	10.2	35.2	76	29	DME546567	AJ546567 Drosophil
c 450	10.2	35.2	77	9	AI955870	AI955870 wt60g03.x
451	10.2	35.2	77	10	BE887372	BE887372 601509504
c 452	10.2	35.2	77	14	CK137902	CK137902 maj98h02.
453	10.2	35.2	77	28	AF067770	AF067770 AF067770
454	10.2	35.2	77	29	AL754972	AL754972 Arabidops
c 455	10.2	35.2	78	10	BG063280	BG063280 H3005C03-
456	10.2	35.2	78	28	BH252197	BH252197 SALK_0129
457	10.2	35.2	78	29	CG629170	CG629170 OST340998
458	10.2	35.2	79	9	AA758378	AA758378 zg45g11.s
c 459	10.2	35.2	79	9	AI132148	AI132148 ub90c05.r
c 460	10.2	35.2	79	9	AA144854	AA144854 mr73c08.r
461	10.2	35.2	79	9	AU260306	AU260306 AU260306
c 462	10.2	35.2	79	14	CA942759	CA942759 ir67a01.x
463	10.2	35.2	79	14	T59637	T59637 yc13c10.s1
c 464	10.2	35.2	80	9	AA066080	AA066080 mm41c02.r
c 465	10.2	35.2	80	9	AJ499317	AJ499317 AJ499317
466	10.2	35.2	80	10	BG021609	BG021609 dg33b06.x
467	10.2	35.2	80	28	AZ808210	AZ808210 2M0071G18
c 468	10.2	35.2	80	29	CC798537	CC798537 SALK_1464
c 469	10.2	35.2	80	29	CG549290	CG549290 OST152453
470	10.2	35.2	80	29	BX659472	BX659472 Arabidops
471	10.2	35.2	80	29	BX894649	BX894649 Arabidops
c 472	10	34.5	31	9	AA928444	AA928444 om77d06.s
473	10	34.5	32	13	C01280	C01280 HUMGS000367
474	10	34.5	33	9	AW059779	AW059779 LE5c02.yg
475	10	34.5	34	28	BH850881	BH850881 SALK_0720
476	10	34.5	35	14	D26041	D26041 HUMGS03671
477	10	34.5	39	9	AU255393	AU255393 AU255393
478	10	34.5	39	28	AZ450474	AZ450474 1M0249A13
479	10	34.5	39	28	BH909815	BH909815 SALK_0561
480	10	34.5	40	29	AL943578	AL943578 Arabidops
c 481	10	34.5	42	28	BZ381341	BZ381341 SALK_1165
482	10	34.5	42	29	BX534624	BX534624 Arabidops
483	10	34.5	43	9	AA096781	AA096781 mo06c06.r
c 484	10	34.5	43	28	BH619407	BH619407 1007059A0
485	10	34.5	45	28	BZ353391	BZ353391 SALK_1202
486	10	34.5	45	28	CC178546	CC178546 NPX482 Ba
487	10	34.5	47	29	TA92D04Q	AL462426 T. brucei
c 488	10	34.5	48	12	BJ014884	BJ014884 BJ014884
489	10	34.5	48	28	BH910024	BH910024 SALK_0572
490	10	34.5	49	29	CC885064	CC885064 SALK_1459
c 491	10	34.5	50	28	AZ510105	AZ510105 1M0354D24
c 492	10	34.5	50	28	BH790272	BH790272 SALK_0567
c 493	10	34.5	51	28	BZ381725	BZ381725 SALK_1171
c 494	10	34.5	51	29	CG696063	CG696063 BARC_BFGL
c 495	10	34.5	51	29	BX126671	BX126671 Danio rer
496	10	34.5	51	29	DME545740	AJ545740 Drosophil
497	10	34.5	52	9	AI180755	AI180755 ub91f10.r
c 498	10	34.5	52	9	AI789498	AI789498 ue80a01.r
c 499	10	34.5	52	10	BF631974	BF631974 NF007E11D
500	10	34.5	52	14	D20654	D20654 HUMGS01629
501	10	34.5	52	28	AZ484394	AZ484394 1M0310M18
c 502	10	34.5	52	28	BZ662046	BZ662046 SALK_0255
503	10	34.5	53	28	BZ353392	BZ353392 SALK_1202
504	10	34.5	53	28	BZ353393	BZ353393 SALK_1202
505	10	34.5	53	28	BZ353394	BZ353394 SALK_1202
c 506	10	34.5	53	28	BZ358385	BZ358385 SALK_1323
507	10	34.5	53	28	BZ382235	BZ382235 SALK_1180
c 508	10	34.5	54	28	AZ665895	AZ665895 1M0547P01
c 509	10	34.5	54	29	BX662547	BX662547 Arabidops
510	10	34.5	55	14	R52718	R52718 yg99b05.r1
511	10	34.5	56	9	AU257644	AU257644 AU257644
512	10	34.5	56	29	AL762058	AL762058 Arabidops
c 513	10	34.5	56	29	BX002727	BX002727 Arabidops
514	10	34.5	57	29	AL760415	AL760415 Arabidops
515	10	34.5	58	9	AI006125	AI006125 ua86h05.r

c 516	10	34.5	58	9	AI824019	AI824019 wj29f03.x
c 517	10	34.5	58	28	AZ649301	AZ649301 1M0518F15
518	10	34.5	58	29	CC569427	CC569427 CH240_444
c 519	10	34.5	59	14	CA794962	CA794962 Cac_BL_20
520	10	34.5	60	9	AA432527	AA432527 vd72h08.r
c 521	10	34.5	60	14	CD393710	CD393710 Gm_ck1328
522	10	34.5	61	9	AI318033	AI318033 ta75g02.x
c 523	10	34.5	61	29	CG513355	CG513355 OST66685
524	10	34.5	63	12	BM300267	BM300267 MCR043C06
525	10	34.5	63	14	CB008549	CB008549 VVC057G04
526	10	34.5	63	14	CF806791	CF806791 psHB016xA
c 527	10	34.5	63	29	CG717239	CG717239 1119047E0
528	10	34.5	63	29	AG021925	AG021925 Oryza sat
529	10	34.5	64	9	AA815034	AA815034 oa88b12.s
c 530	10	34.5	64	9	AA908483	AA908483 og82d01.s
531	10	34.5	64	9	AA939303	AA939303 ol78g12.s
c 532	10	34.5	64	9	AA102414	AA102414 zm26a04.s
533	10	34.5	64	9	AA576602	AA576602 nm66h10.s
534	10	34.5	64	14	CD940615	CD940615 RAP_4 Gen
535	10	34.5	64	14	CD955522	CD955522 SBW_124 G
c 536	10	34.5	64	29	CG718710	CG718710 1119054B0
c 537	10	34.5	64	29	CG718764	CG718764 1119054D0
538	10	34.5	65	9	AA691167	AA691167 vt34a01.r
539	10	34.5	65	12	BM530898	BM530898 fw97b06.x
c 540	10	34.5	65	29	AL768493	AL768493 Arabidops
c 541	10	34.5	66	12	BM569368	BM569368 kj60c04.y
542	10	34.5	66	28	BH759565	BH759565 KG05089-3
543	10	34.5	66	29	BX531394	BX531394 Arabidops
544	10	34.5	67	9	AI905592	AI905592 CM-BT094-
545	10	34.5	68	29	TA152H06P	AL467222 T. brucei
546	10	34.5	69	9	AA213331	AA213331 mw83h08.r
547	10	34.5	69	9	AV535100	AV535100 AV535100
c 548	10	34.5	69	10	BE491970	BE491970 GREB138 e
549	10	34.5	69	28	AZ620587	AZ620587 1M0453F19
c 550	10	34.5	69	28	BH792044	BH792044 SALK_0625
c 551	10	34.5	69	29	CC883497	CC883497 SALK_0944
552	10	34.5	69	29	CG504738	CG504738 OST53105
553	10	34.5	70	9	AI824448	AI824448 tx70d03.x
c 554	10	34.5	70	9	AA516989	AA516989 vh89d02.r
c 555	10	34.5	70	13	BQ275042	BQ275042 pj40e06.y
556	10	34.5	70	14	T97779	T97779 ye58h04.s1
557	10	34.5	70	29	CG545898	CG545898 OST144669
c 558	10	34.5	70	29	BX892348	BX892348 Arabidops
559	10	34.5	71	12	BM873635	BM873635 laa01a12.
560	10	34.5	71	14	CD955658	CD955658 SBW_68 Ge
561	10	34.5	71	28	AZ833202	AZ833202 2M0115E08
562	10	34.5	72	28	BH852570	BH852570 SALK_0752
c 563	10	34.5	72	28	BZ596806	BZ596806 SALK_0960
c 564	10	34.5	72	29	CG722933	CG722933 1119074A1
c 565	10	34.5	73	9	AA932772	AA932772 om55h12.s
566	10	34.5	73	9	AA097948	AA097948 mn83b11.r
567	10	34.5	73	9	AI744222	AI744222 tr07f10.x
c 568	10	34.5	73	12	BM889450	BM889450 laa01a12.
569	10	34.5	73	28	AZ391057	AZ391057 1M0152F22
c 570	10	34.5	73	28	BH856093	BH856093 SALK_0834
571	10	34.5	73	29	CG483091	CG483091 OST16553
c 572	10	34.5	73	29	CG529491	CG529491 OST109956
c 573	10	34.5	73	29	BX291303	BX291303 Arabidops
574	10	34.5	73	29	CNS0428W	AL271193 Tetraodon
575	10	34.5	75	28	B46656	B46656 HS-1065-B1-
576	10	34.5	75	29	AL771768	AL771768 Arabidops
c 577	10	34.5	76	9	AA936354	AA936354 on71c08.s
578	10	34.5	76	9	AI708190	AI708190 as89g08.x
579	10	34.5	76	13	BQ536895	BQ536895 STEM2_18_
580	10	34.5	76	28	AZ622889	AZ622889 1M0460B02
581	10	34.5	76	28	AZ657549	AZ657549 1M0533L18
c 582	10	34.5	76	29	CG658835	CG658835 OST435402
583	10	34.5	76	29	BX292165	BX292165 Arabidops
c 584	10	34.5	77	10	AW700694	AW700694 pa43b07.y
585	10	34.5	77	13	BU287571	BU287571 604166411
586	10	34.5	77	29	CG558425	CG558425 OST175926
587	10	34.5	78	10	BG059210	BG059210 nah51f12.
c 588	10	34.5	78	28	AZ594192	AZ594192 1M0406H15
c 589	10	34.5	78	29	AL756665	AL756665 Arabidops
c 590	10	34.5	78	29	AL953578	AL953578 Arabidops
c 591	10	34.5	79	9	AI318547	AI318547 ta74h10.x
592	10	34.5	79	9	AU268041	AU268041 AU268041
593	10	34.5	79	12	BG363804	BG363804 dc71e03.y
594	10	34.5	79	28	AZ783177	AZ783177 2M0024G10
595	10	34.5	79	28	BH902542	BH902542 SALK_0919
c 596	10	34.5	79	28	BZ594513	BZ594513 SALK_0841
597	10	34.5	79	29	CG520057	CG520057 OST84820
598	10	34.5	79	29	CG617952	CG617952 OST312048
599	10	34.5	79	29	AL760717	AL760717 Arabidops
600	10	34.5	80	9	AA691168	AA691168 vt34a02.r
601	10	34.5	80	28	AZ345097	AZ345097 1M0079C01
c 602	10	34.5	80	28	BH865294	BH865294 SALK_0981
c 603	9.8	33.8	27	2	HSM003610	Al039134 Homo sapi
604	9.8	33.8	27	14	D21047	D21047 HUMGS02032
c 605	9.8	33.8	29	28	AZ810362	AZ810362 2M0074P16
c 606	9.8	33.8	30	2	HSM001042	Al036726 Homo sapi
607	9.8	33.8	30	2	HSM001581	Al037255 Homo sapi
c 608	9.8	33.8	31	2	HSM002270	Al037936 Homo sapi
609	9.8	33.8	32	2	HSM001595	Al037269 Homo sapi
610	9.8	33.8	32	2	HSM001674	Al037346 Homo sapi
c 611	9.8	33.8	33	2	HSM001806	Al037475 Homo sapi
612	9.8	33.8	33	2	HSM003086	Al038610 Homo sapi
613	9.8	33.8	33	2	HSM003165	Al038689 Homo sapi
c 614	9.8	33.8	33	9	AL048719	AL048719 DKFZp566G
c 615	9.8	33.8	33	9	AL048733	AL048733 DKFZp566I
616	9.8	33.8	34	2	HSM003069	Al038593 Homo sapi
c 617	9.8	33.8	34	9	AI744308	AI744308 tr09b07.x
c 618	9.8	33.8	35	12	BJ016002	BJ016002 BJ016002
c 619	9.8	33.8	36	29	BX890659	BX890659 Arabidops
c 620	9.8	33.8	37	9	AI015197	AI015197 ot82g11.s
621	9.8	33.8	37	9	AI802260	AI802260 tj36g07.x
622	9.8	33.8	38	28	AZ500985	AZ500985 1M0339F07
c 623	9.8	33.8	40	2	HSM001818	Al037487 Homo sapi
c 624	9.8	33.8	40	28	AZ510911	AZ510911 1M0355N08
c 625	9.8	33.8	41	29	BX891103	BX891103 Arabidops
c 626	9.8	33.8	42	10	BF133132	BF133132 601645566
627	9.8	33.8	42	14	CA797521	CA797521 Cac_BL_46
c 628	9.8	33.8	42	14	CF306863	CF306863 HDAI--05-
c 629	9.8	33.8	42	29	CC792858	CC792858 SALK_0025
c 630	9.8	33.8	43	9	AA744540	AA744540 ny79b09.s
631	9.8	33.8	44	28	AZ873907	AZ873907 2M0187A20
632	9.8	33.8	44	29	AL771069	AL771069 Arabidops
633	9.8	33.8	44	29	BX172399	BX172399 Danio rer
634	9.8	33.8	45	29	AL765541	AL765541 Arabidops
635	9.8	33.8	46	9	AA622161	AA622161 nq56f10.s
636	9.8	33.8	46	29	AL951447	AL951447 Arabidops
637	9.8	33.8	47	9	AA004970	AA004970 zh94b03.r
638	9.8	33.8	47	14	CF305286	CF305286 CLD1--01-
639	9.8	33.8	47	28	AZ603333	AZ603333 1M0422D14
c 640	9.8	33.8	47	28	AZ766816	AZ766816 1M0564E10
c 641	9.8	33.8	47	29	TA121B05Q	AL463003 T. brucei
642	9.8	33.8	48	28	AZ317194	AZ317194 1M0035I12
643	9.8	33.8	48	28	BZ763599	BZ763599 SALK_1195
644	9.8	33.8	48	29	AL760281	AL760281 Arabidops
c 645	9.8	33.8	49	9	AA648244	AA648244 ns07h03.r
646	9.8	33.8	49	9	AV962538	AV962538 AV962538
c 647	9.8	33.8	49	14	W81217	W81217 zd85d07.s1
c 648	9.8	33.8	49	14	W86814	W86814 zh61f02.r1
c 649	9.8	33.8	49	28	AQ025156	AQ025156 EP(3)1038
650	9.8	33.8	49	28	AZ576537	AZ576537 AST-T11C0
c 651	9.8	33.8	49	28	CC179048	CC179048 SALK_0576
c 652	9.8	33.8	50	9	AU104262	AU104262 AU104262
653	9.8	33.8	50	9	AU104387	AU104387 AU104387
c 654	9.8	33.8	50	28	AQ026189	AQ026189 l(3)03806
c 655	9.8	33.8	51	13	BQ593814	BQ593814 E012765-0
c 656	9.8	33.8	51	14	CD682301	CD682301 rj49d02.y
c 657	9.8	33.8	51	29	HSMC48G03	X88005 H.sapiens D
c 658	9.8	33.8	51	29	TA289F12Q	AL486591 T. brucei
c 659	9.8	33.8	52	9	AA575934	AA575934 nm56c11.s
660	9.8	33.8	52	10	BF634731	BF634731 NF068D05D
661	9.8	33.8	52	10	BF635399	BF635399 NF068D04D

	662	9.8	33.8	52	10	AW245287 AW245287 2820140.3
c	663	9.8	33.8	52	14	CD683281 CD683281 rj41f08.y
	664	9.8	33.8	52	29	BX892212 BX892212 Arabidops
	665	9.8	33.8	53	2	HSM002757 Al038411 Homo sapi
	666	9.8	33.8	53	12	BG152242 BG152242 nag74f05.
	667	9.8	33.8	53	28	AZ467391 AZ467391 1M0278E18
	668	9.8	33.8	53	29	AL942442 AL942442 Arabidops
	669	9.8	33.8	54	12	BG917121 BG917121 602816448
	670	9.8	33.8	54	13	BQ537862 BQ537862 pz06b07.y
c	671	9.8	33.8	54	28	AZ607311 AZ607311 1M0429M12
c	672	9.8	33.8	54	28	AZ783536 AZ783536 2M0025F23
	673	9.8	33.8	54	29	CC889004 CC889004 SALK_1526
	674	9.8	33.8	56	14	H09345 H09345 yl95g01.s1
c	675	9.8	33.8	56	14	T75345 T75345 yc90c10.r1
c	676	9.8	33.8	56	28	AZ785122 AZ785122 2M0028H14
c	677	9.8	33.8	56	28	B04383 B04383 cSRL-31c7-u
	678	9.8	33.8	57	12	BI492632 BI492632 df27a08.w
c	679	9.8	33.8	57	28	BH849364 BH849364 SALK_0695
c	680	9.8	33.8	57	28	BH850801 BH850801 SALK_0718
	681	9.8	33.8	57	29	CC794042 CC794042 SALK_0405
c	682	9.8	33.8	58	9	AA732710 AA732710 nz85e08.s
c	683	9.8	33.8	58	9	AI965857 AI965857 sc78g02.y
	684	9.8	33.8	58	9	AA480462 AA480462 ne70d10.s
c	685	9.8	33.8	58	12	BJ064519 BJ064519
	686	9.8	33.8	58	14	CA953677 CA953677 pt68e10.y
	687	9.8	33.8	58	14	CA953683 CA953683 pt68g04.y
c	688	9.8	33.8	58	28	BZ378649 BZ378649 SALK_1084
	689	9.8	33.8	59	2	HSM001952 Al037621 Homo sapi
c	690	9.8	33.8	59	9	AA905170 AA905170 ok06c11.s
c	691	9.8	33.8	59	9	AA620542 AA620542 ae61d10.s
	692	9.8	33.8	59	12	BM285373 BM285373 EST00014
	693	9.8	33.8	59	12	BM493229 BM493229 EST00012
	694	9.8	33.8	59	13	BQ094074 BQ094074 040802_33
c	695	9.8	33.8	59	13	BQ094087 BQ094087 040802_46
c	696	9.8	33.8	59	28	B00362 B00362 cSRL-108d12
	697	9.8	33.8	59	28	BZ354569 BZ354569 SALK_1254
	698	9.8	33.8	60	10	BE871815 BE871815 601447803
	699	9.8	33.8	60	14	R36900 R36900 yf63a09.s1
	700	9.8	33.8	61	9	AA715903 AA715903 nv76d06.r
c	701	9.8	33.8	61	9	AL774563 AL774563 AL774563
	702	9.8	33.8	61	9	AA413732 AA413732 vc55h04.s
c	703	9.8	33.8	61	9	AA587759 AA587759 nm95e11.s
c	704	9.8	33.8	61	10	BE740294 BE740294 601595279
	705	9.8	33.8	61	14	CB366487 CB366487 ZF001-P00
c	706	9.8	33.8	61	14	CD962460 CD962460 SDP_160 G
c	707	9.8	33.8	62	14	CD953331 CD953331 SBH_233 G
	708	9.8	33.8	62	14	CF543050 CF543050 S015532-0
c	709	9.8	33.8	62	14	R46765 R46765 yj54e01.s1
	710	9.8	33.8	62	29	DME547034 AJ547034 Drosophil
	711	9.8	33.8	63	14	CD945613 CD945613 REC_12 Ge
	712	9.8	33.8	63	14	CD967872 CD967872 SEY_34 Ge
	713	9.8	33.8	63	14	CF421525 CF421525 jaa11c10.
	714	9.8	33.8	63	28	AZ820521 AZ820521 2M0092D10
c	715	9.8	33.8	63	29	CG983379 CG983379 CH240_165
c	716	9.8	33.8	64	9	AA724842 AA724842 ah97b04.s
c	717	9.8	33.8	64	9	AI254470 AI254470 qv47h04.x
	718	9.8	33.8	64	10	AW626420 AW626420 SWOvAFCAP
	719	9.8	33.8	64	10	BE636255 BE636255 SWOvAMCAQ
c	720	9.8	33.8	64	28	AZ775161 AZ775161 2M0007F15
	721	9.8	33.8	64	28	BH405785 BH405785 RPCI-23-2
c	722	9.8	33.8	64	29	AL763069 AL763069 Arabidops
c	723	9.8	33.8	65	2	HSM001786 Al037455 Homo sapi
	724	9.8	33.8	65	13	BX735697 BX735697 BX735697
c	725	9.8	33.8	65	14	CD391037 CD391037 Gm_ck0466
c	726	9.8	33.8	65	28	BH856124 BH856124 SALK_0833
c	727	9.8	33.8	65	29	CG774219 CG774219 1123018A0
	728	9.8	33.8	66	28	BH811735 BH811735 SALK_0596
c	729	9.8	33.8	67	9	AA778316 AA778316 zf35d06.s
c	730	9.8	33.8	67	9	AI271601 AI271601 qw84f10.x
c	731	9.8	33.8	67	9	AA506572 AA506572 ni17d07.s
c	732	9.8	33.8	67	14	CB275259 CB275259 ku56e02.y
c	733	9.8	33.8	67	14	CB357125 CB357125 ZF001-P00
	734	9.8	33.8	67	14	CD953649 CD953649 SBJ_4 Gen
	735	9.8	33.8	67	28	AZ621034 AZ621034 1M0454F03
c	736	9.8	33.8	67	29	CG620673 CG620673 OST318050
	737	9.8	33.8	68	9	AL871272 AL871272 AL871272
	738	9.8	33.8	68	12	BI155193 BI155193 602903173
	739	9.8	33.8	68	28	AZ803353 AZ803353 2M0063N11
	740	9.8	33.8	69	13	BQ126724 BQ126724 ii18e01.y
	741	9.8	33.8	69	14	CA840803 CA840803 pt59a03.y
	742	9.8	33.8	69	14	CD843726 CD843726 RFO2.133L
c	743	9.8	33.8	69	29	CG532276 CG532276 OST115959
c	744	9.8	33.8	69	29	CG730488 CG730488 1119127D0
	745	9.8	33.8	70	9	AI028481 AI028481 oy74a11.x
c	746	9.8	33.8	70	28	AY127152 AY127152
	747	9.8	33.8	70	29	AL758061 AL758061 Arabidops
c	748	9.8	33.8	71	13	BQ548401 BQ548401 rd25h07.y
c	749	9.8	33.8	71	14	CF269849 CF269849 Fcylcold1
	750	9.8	33.8	71	14	CF808257 CF808257 psHB033xI
	751	9.8	33.8	71	28	AZ786367 AZ786367 2M0031O16
c	752	9.8	33.8	71	28	AZ804199 AZ804199 2M0065E02
c	753	9.8	33.8	71	29	CG520678 CG520678 OST86569
c	754	9.8	33.8	72	12	BI437885 BI437885 ic81a10.x
	755	9.8	33.8	72	13	C01172 C01172 HUMGS000786
	756	9.8	33.8	72	14	CD964236 CD964236 SEA_82 Ge
c	757	9.8	33.8	72	14	X85577 X85577 HS150EST h
	758	9.8	33.8	72	28	BH846818 BH846818 SALK_0105
c	759	9.8	33.8	72	28	BZ661345 BZ661345 SALK_0248
	760	9.8	33.8	72	29	CG590866 CG590866 OST244762
	761	9.8	33.8	72	29	BX286715 BX286715 Arabidops
c	762	9.8	33.8	72	29	BX653165 BX653165 Arabidops
	763	9.8	33.8	73	9	AI856019 AI856019 sc30c11.x
	764	9.8	33.8	73	9	AL693794 AL693794 AL693794
c	765	9.8	33.8	73	9	AV532998 AV532998 AV532998
	766	9.8	33.8	73	13	BU815920 BU815920 N048G03 P
c	767	9.8	33.8	73	28	AZ453813 AZ453813 1M0255N22
c	768	9.8	33.8	73	28	AZ797140 AZ797140 2M0053M10
	769	9.8	33.8	73	28	AZ805957 AZ805957 2M0067I10
	770	9.8	33.8	73	28	BH865938 BH865938 SALK_1001
c	771	9.8	33.8	73	29	CG511759 CG511759 OST64281
c	772	9.8	33.8	73	29	AL764709 AL764709 Arabidops
	773	9.8	33.8	73	29	BX657393 BX657393 Arabidops
c	774	9.8	33.8	74	9	AI819570 AI819570 wj91g08.x
	775	9.8	33.8	74	14	CA953696 CA953696 pt69d05.y
	776	9.8	33.8	74	28	BZ595894 BZ595894 SALK_0897
	777	9.8	33.8	74	29	CC797033 CC797033 SALK_1443
	778	9.8	33.8	75	14	CA953605 CA953605 pt66a12.y
	779	9.8	33.8	75	28	BZ382121 BZ382121 SALK_1178
c	780	9.8	33.8	75	29	CG644720 CG644720 OST388416
	781	9.8	33.8	75	29	BX290729 BX290729 Arabidops
c	782	9.8	33.8	76	9	AA975663 AA975663 oq63f07.s
c	783	9.8	33.8	76	12	BM354196 BM354196 if31f01.y
	784	9.8	33.8	76	13	BQ818585 BQ818585 1030071H0
	785	9.8	33.8	76	13	BU652815 BU652815 1112102E1
	786	9.8	33.8	76	14	CA840682 CA840682 pt54b05.y
c	787	9.8	33.8	76	28	CC156059 CC156059 NPX098 Ba
c	788	9.8	33.8	77	9	AA733450 AA733450 vt74g01.r
	789	9.8	33.8	77	13	BQ456810 BQ456810 ke31c12.y
c	790	9.8	33.8	77	13	BQ821575 BQ821575 1030093D0
c	791	9.8	33.8	77	14	CF115543 CF115543 SSA005N05
	792	9.8	33.8	77	14	R32307 R32307 yh68b03.s1
c	793	9.8	33.8	77	28	BZ380165 BZ380165 SALK_1147
	794	9.8	33.8	77	29	BX290041 BX290041 Arabidops
	795	9.8	33.8	78	14	CD885125 CD885125 G100D.001
	796	9.8	33.8	78	28	AZ919077 AZ919077 1006013H0
c	797	9.8	33.8	78	28	BH855366 BH855366 SALK_0859
c	798	9.8	33.8	78	28	BZ662137 BZ662137 SALK_0256
c	799	9.8	33.8	78	28	CC326468 CC326468 XN369 Bay
c	800	9.8	33.8	79	9	AA782367 AA782367 ac24e09.s
c	801	9.8	33.8	79	9	AV955805 AV955805 AV955805
	802	9.8	33.8	79	13	BU872819 BU872819 Q047C09 P
c	803	9.8	33.8	79	13	BX556305 BX556305
	804	9.8	33.8	79	29	CG427084 CG427084 01S0723-0
	805	9.8	33.8	79	29	CG538124 CG538124 OST127760
	806	9.8	33.8	79	29	AL764001 AL764001 Arabidops
	807	9.8	33.8	79	29	AL946773 AL946773 Arabidops

c 808	9.8	33.8	79	29	BX004713	BX004713 Arabidops
809	9.8	33.8	80	9	AI215932	AI215932 qm41d07.x
c 810	9.8	33.8	80	12	BI708952	BI708952 fp94f11.y
c 811	9.8	33.8	80	13	BX253137	BX253137
c 812	9.8	33.8	80	14	CD395443	CD395443 Gm_ck1546
813	9.8	33.8	80	14	CK114217	CK114217 V054D03 P
814	9.8	33.8	80	14	N55635	N55635 ESTG183 Rat
c 815	9.8	33.8	80	28	AZ566704	AZ566704 225PvA08
816	9.8	33.8	80	28	BH910286	BH910286 SALK_0587
817	9.8	33.8	80	29	CG427082	CG427082 01S0723-0
c 818	9.8	33.8	80	29	CG542129	CG542129 OST136154
819	9.6	33.1	21	28	AZ832301	AZ832301 2M0112F10
820	9.6	33.1	30	29	TA49E08P	AL456886 T. brucei
c 821	9.6	33.1	34	9	AA931137	AA931137 oo70b07.s
822	9.6	33.1	34	9	AU258702	AU258702 AU258702
c 823	9.6	33.1	37	28	AZ308218	AZ308218 1M0011D02
824	9.6	33.1	37	29	AL759196	AL759196 Arabidops
825	9.6	33.1	38	29	AG205629	AG205629 Oryza sat
c 826	9.6	33.1	39	28	AZ312577	AZ312577 1M0028D06
827	9.6	33.1	39	28	AZ774271	AZ774271 2M0003A03
c 828	9.6	33.1	39	28	BZ770677	BZ770677 SALK_1436
829	9.6	33.1	40	9	AA933656	AA933656 om56d08.s
830	9.6	33.1	40	9	AA112852	AA112852 zm63c10.s
831	9.6	33.1	40	29	CG778161	CG778161 1123026E0
832	9.6	33.1	41	14	CF973215	CF973215 PSU_bifou
833	9.6	33.1	41	28	BZ383963	BZ383963 SALK_1348
834	9.6	33.1	42	29	CC889080	CC889080 SALK_1527
835	9.6	33.1	43	13	BQ587208	BQ587208 E012349-0
c 836	9.6	33.1	43	28	AZ798954	AZ798954 2M0056J06
837	9.6	33.1	44	28	BH808431	BH808431 1008079G1
c 838	9.6	33.1	45	9	AU265554	AU265554 AU265554
c 839	9.6	33.1	46	29	TA363H04P	AL494300 T. brucei
840	9.6	33.1	47	28	BH891916	BH891916 3526_1_19
841	9.6	33.1	48	9	AV956088	AV956088 AV956088
842	9.6	33.1	48	12	BI459090	BI459090 603199445
843	9.6	33.1	48	28	BH846418	BH846418 SALK_0078
844	9.6	33.1	48	28	BZ291472	BZ291472 SALK_1207
c 845	9.6	33.1	48	29	CG718613	CG718613 1119053G0
846	9.6	33.1	48	29	CG780232	CG780232 1123038C0
c 847	9.6	33.1	48	29	AL945380	AL945380 Arabidops
848	9.6	33.1	49	14	CF321218	CF321218 HD--12-G0
849	9.6	33.1	50	9	AU102950	AU102950 AU102950
850	9.6	33.1	50	9	AU103049	AU103049 AU103049
851	9.6	33.1	50	9	AU103051	AU103051 AU103051
852	9.6	33.1	50	9	AU105371	AU105371 AU105371
853	9.6	33.1	50	29	AL945912	AL945912 Arabidops
c 854	9.6	33.1	50	29	BX285477	BX285477 Arabidops
855	9.6	33.1	50	29	BX655656	BX655656 Arabidops
c 856	9.6	33.1	51	9	AV741278	AV741278 AV741278
c 857	9.6	33.1	51	29	CG717347	CG717347 1119048B0
c 858	9.6	33.1	52	9	AA102545	AA102545 zn41f01.s
c 859	9.6	33.1	52	9	AI697308	AI697308 tq07f02.x
c 860	9.6	33.1	52	9	AL868703	AL868703 AL868703
861	9.6	33.1	52	9	AU014097	AU014097 AU014097
c 862	9.6	33.1	52	14	CD682289	CD682289 rj49b06.y
863	9.6	33.1	52	28	AZ843322	AZ843322 2M0142G09
864	9.6	33.1	53	9	AU259280	AU259280 AU259280
c 865	9.6	33.1	53	9	AA529390	AA529390 vi39g06.r
c 866	9.6	33.1	53	28	AZ859610	AZ859610 2M0165I15
867	9.6	33.1	53	28	CC457487	CC457487 SALK_1102
c 868	9.6	33.1	54	9	AU257589	AU257589 AU257589
869	9.6	33.1	54	12	BG272439	BG272439 nah30f09.
c 870	9.6	33.1	54	29	CC888314	CC888314 SALK_1516
c 871	9.6	33.1	54	29	AL755363	AL755363 Arabidops
c 872	9.6	33.1	55	9	AA886619	AA886619 ny42e03.s
c 873	9.6	33.1	55	9	AI913453	AI913453 tz77e09.x
874	9.6	33.1	55	28	BH629048	BH629048 1007076B0
875	9.6	33.1	55	28	BH911635	BH911635 SALK_0698
876	9.6	33.1	55	28	BZ377638	BZ377638 SALK_0837
877	9.6	33.1	55	28	BZ664325	BZ664325 SALK_0698
c 878	9.6	33.1	55	29	BX894815	BX894815 Arabidops
879	9.6	33.1	56	13	BQ787611	BQ787611 im13b12.x
880	9.6	33.1	56	28	AZ363074	AZ363074 1M0108L24
c 881	9.6	33.1	56	28	AZ424951	AZ424951 1M0204K15
882	9.6	33.1	56	28	AZ467755	AZ467755 1M0279C22
883	9.6	33.1	56	28	BZ665747	BZ665747 KG10262 D
c 884	9.6	33.1	57	12	BI156096	BI156096 602903657
885	9.6	33.1	57	13	BX553132	BX553132 BX553132
886	9.6	33.1	57	28	BH410651	BH410651 1007019D1
c 887	9.6	33.1	57	28	BH905116	BH905116 SALK_1056
c 888	9.6	33.1	57	28	BH905124	BH905124 SALK_1056
889	9.6	33.1	57	29	AL758068	AL758068 Arabidops
c 890	9.6	33.1	57	29	TA29E05Q	AL452663 T. brucei
c 891	9.6	33.1	58	9	AA659472	AA659472 nu24g09.s
892	9.6	33.1	58	9	AI052132	AI052132 oy30c12.x
893	9.6	33.1	58	12	BJ080589	BJ080589 BJ080589
894	9.6	33.1	58	12	BM092179	BM092179 sah08h05.
895	9.6	33.1	58	12	BM092295	BM092295 sah11a05.
c 896	9.6	33.1	58	28	B02943	B02943 cSRL-163G2-
897	9.6	33.1	58	29	AL942674	AL942674 Arabidops
898	9.6	33.1	59	9	AV852079	AV852079 AV852079
c 899	9.6	33.1	59	10	AW412926	AW412926 uq50h01.y
900	9.6	33.1	60	9	AV531500	AV531500 AV531500
c 901	9.6	33.1	60	10	BE323490	BE323490 NF009C01P
902	9.6	33.1	60	12	BI864799	BI864799 ft17a05.x
903	9.6	33.1	60	12	BI864831	BI864831 ft18a04.x
904	9.6	33.1	60	12	BJ076533	BJ076533 BJ076533
c 905	9.6	33.1	60	14	CD949567	CD949567 SAL_68 Ge
906	9.6	33.1	60	28	BH856804	BH856804 SALK_0791
c 907	9.6	33.1	61	9	AI906736	AI906736 QV-BT124-
908	9.6	33.1	61	10	AW432632	AW432632 sh83c10.y
c 909	9.6	33.1	61	12	BM123335	BM123335 L0523C08-
910	9.6	33.1	61	28	CC040535	CC040535 3591_1_13
c 911	9.6	33.1	61	29	CG513867	CG513867 OST67472
912	9.6	33.1	61	29	CG534150	CG534150 OST120854
913	9.6	33.1	61	29	CG596505	CG596505 OST258430
914	9.6	33.1	61	29	CG653174	CG653174 OST418177
915	9.6	33.1	61	29	AG224181	AG224181 Lotus cor
c 916	9.6	33.1	61	29	TA215D12Q	AL479253 T. brucei
917	9.6	33.1	62	10	BF638398	BF638398 NF057C04P
918	9.6	33.1	62	10	BF701269	BF701269 602128354
c 919	9.6	33.1	62	12	BM343010	BM343010 fw50e06.y
920	9.6	33.1	62	13	BQ482579	BQ482579 ke51c11.y
c 921	9.6	33.1	62	13	BQ740728	BQ740728 saq51c12.
922	9.6	33.1	62	13	C21188	C21188 HUMGS000220
c 923	9.6	33.1	62	14	CD940564	CD940564 RAO_54 Ge
c 924	9.6	33.1	62	14	H75004	H75004 670 Random-
925	9.6	33.1	62	28	BH855036	BH855036 SALK_0867
926	9.6	33.1	62	29	CG493795	CG493795 OST32270
c 927	9.6	33.1	62	29	CG506365	CG506365 OST55868
c 928	9.6	33.1	62	29	CG541791	CG541791 OST135391
929	9.6	33.1	62	29	CG642006	CG642006 OST377531
c 930	9.6	33.1	62	29	CG646460	CG646460 OST392820
931	9.6	33.1	62	29	AL762490	AL762490 Arabidops
c 932	9.6	33.1	63	12	BG362361	BG362361 gb70h04.y
c 933	9.6	33.1	63	28	AZ478246	AZ478246 1M0298K11
934	9.6	33.1	63	28	BH900972	BH900972 KG07915-5
c 935	9.6	33.1	63	28	BZ597224	BZ597224 SALK_1005
936	9.6	33.1	63	29	CC796002	CC796002 SALK_0918
937	9.6	33.1	63	29	CC796134	CC796134 SALK_0930
938	9.6	33.1	63	29	CG562609	CG562609 OST185202
c 939	9.6	33.1	63	29	CG627471	CG627471 OST337546
940	9.6	33.1	63	29	CG627573	CG627573 OST337884
c 941	9.6	33.1	63	29	CG670922	CG670922 OST471885
942	9.6	33.1	63	29	BX530833	BX530833 Arabidops
943	9.6	33.1	64	9	AI241145	AI241145 qk05e08.x
c 944	9.6	33.1	64	9	AA107029	AA107029 ml92e09.r
945	9.6	33.1	64	9	AI670974	AI670974 wb12f03.x
946	9.6	33.1	64	9	AI800789	AI800789 wg13b07.x
c 947	9.6	33.1	64	10	BF118530	BF118530 SWOvL3CAN
c 948	9.6	33.1	64	13	BQ564870	BQ564870 gi26d01.y
c 949	9.6	33.1	64	28	AZ801237	AZ801237 2M0059D07
950	9.6	33.1	64	28	BH814721	BH814721 SALK_0668
c 951	9.6	33.1	64	28	BH857811	BH857811 SALK_0874
c 952	9.6	33.1	64	29	CG527719	CG527719 OST105952
c 953	9.6	33.1	64	29	CG549567	CG549567 OST153388

	954	9.6	33.1	64	29	CG601792 CG601792 OST274498
c	955	9.6	33.1	64	29	BX894767 BX894767 Arabidops
c	956	9.6	33.1	65	9	AI279493 AI279493 ql51a01.x
c	957	9.6	33.1	65	13	BX744236 BX744236
c	958	9.6	33.1	65	14	CD943273 CD943273 RCN_128 G
	959	9.6	33.1	65	14	CD943324 CD943324 RCO_26 Ge
c	960	9.6	33.1	65	14	CD963195 CD963195 SDT_105 G
	961	9.6	33.1	65	14	D12425 D12425 HUM00TW151
c	962	9.6	33.1	65	28	AZ789730 AZ789730 2M0037I17
c	963	9.6	33.1	65	28	BH754835 BH754835 SALK_0442
	964	9.6	33.1	65	28	BZ358754 BZ358754 SALK_1332
c	965	9.6	33.1	65	29	CG475746 CG475746 OST4309 M
c	966	9.6	33.1	65	29	CG566789 CG566789 OST192684
c	967	9.6	33.1	65	29	CG601056 CG601056 OST271968
c	968	9.6	33.1	65	29	CG664319 CG664319 OST451176
	969	9.6	33.1	65	29	FR0053439 AL689718 Fugu rubr
c	970	9.6	33.1	66	9	AW041991 AW041991 SWYACAL02
	971	9.6	33.1	66	14	CB051414 CB051414 NISC_gj24
	972	9.6	33.1	66	28	AZ978729 AZ978729 2M0255L09
c	973	9.6	33.1	66	29	CG549359 CG549359 OST152667
c	974	9.6	33.1	66	29	CG606291 CG606291 OST284006
	975	9.6	33.1	67	9	AA894821 AA894821 oj61d12.s
	976	9.6	33.1	67	9	AI365188 AI365188 qx97e01.x
c	977	9.6	33.1	67	10	BE027305 BE027305 EtESTea89
	978	9.6	33.1	67	13	BX707598 BX707598 BX707598
	979	9.6	33.1	67	14	CB261630 CB261630 93-E8866-
c	980	9.6	33.1	67	14	CF776278 CF776278 jaa17f08.
	981	9.6	33.1	67	14	H26340 H26340 yl55d07.r1
	982	9.6	33.1	67	28	BH227336 BH227336 1006139C0
	983	9.6	33.1	67	28	BZ766523 BZ766523 SALK_1375
c	984	9.6	33.1	67	29	BX120588 BX120588 Danio rer
c	985	9.6	33.1	67	29	BX204575 BX204575 Danio rer
	986	9.6	33.1	67	29	BX288073 BX288073 Arabidops
	987	9.6	33.1	67	29	TA272C01P AL483915 T. brucei
	988	9.6	33.1	68	9	AA710447 AA710447 vt42f02.r
c	989	9.6	33.1	68	9	AA545591 AA545591 vj90h07.s
c	990	9.6	33.1	68	9	AA604953 AA604953 no93a10.s
c	991	9.6	33.1	68	12	BG523777 BG523777 35-44 Ste
	992	9.6	33.1	68	13	BQ586079 BQ586079 E012394-0
c	993	9.6	33.1	68	14	CD682098 CD682098 rj46c08.y
c	994	9.6	33.1	68	28	AZ637209 AZ637209 1M0496P24
c	995	9.6	33.1	68	28	BH803478 BH803478 1008100G1
c	996	9.6	33.1	68	29	CC792773 CC792773 SALK_0007
c	997	9.6	33.1	68	29	CG550201 CG550201 OST155596
c	998	9.6	33.1	68	29	CG644709 CG644709 OST388378
	999	9.6	33.1	68	29	BX656783 BX656783 Arabidops
	1000	9.6	33.1	69	9	AA975435 AA975435 oq28h08.s

ALIGNMENTS

RESULT 1
AZ833686
LOCUS AZ833686 46 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0115L20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0115L20 R, genomic survey sequence.
ACCESSION AZ833686
VERSION AZ833686.1 GI:13003594
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0115 row: L column: 20
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 46.
FEATURES Location/Qualifiers
source 1. .46
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0115L20"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
ORIGIN

Query Match 44.8%; Score 13; DB 28; Length 46;
Best Local Similarity 50.0%; Pred. No. 3.1e+04;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGC 19
              ||: |::: |:||||
Db          8 GATACTTTAAGTAAGC 23
```

RESULT 2
BH759592
LOCUS BH759592 70 bp DNA linear GSS 12-MAR-2002
DEFINITION KG05236-3prime Drosophila melanogaster P{SUPor-P} P element insertion lines Drosophila melanogaster genomic Sequence recovered from 3' end of P element, genomic survey sequence.
ACCESSION BH759592
VERSION BH759592.1 GI:19352831
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 70)
AUTHORS Levis,R., Hoskins,R., Liao,G., Mozden,N., Tsang,G., He,Y., Karpen,G., Bellen,H., Rubin,G. and Spradling,A.
TITLE The Berkeley Drosophila Genome Project Gene Disruption Project
JOURNAL Unpublished (2001)
COMMENT Contact: Gerald Rubin
Berkeley Drosophila Genome Project

```
           University of California, Berkeley
           LSA Building, Berkeley, CA 94720-3200, USA
           Fax: 5106439947
           Email: gerry@fruitfly.berkeley.edu
           Sequence recovery method was inverse PCR.
            Sequence orientation is forward strand relative to 5' end of P
           element
            The P element insertion position is base 1 in the 70 bases. This
           insertion position refers to the first base of the 8 base target
           recognition sequence.
           Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster P{SUPor-P} P element
                     insertion lines"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains one or more
                     P{SUPor-P} P-element transposon insertion. The resultant
                     fragment for each strain was directly sequenced to
                     determine the genomic sequence at the site of insertion.
                     Details of the protocols used can be found at
                     http://www.fruitfly.org/about/methods/inverse.pcr.html."
ORIGIN

  Query Match             44.1%;  Score 12.8;  DB 28;  Length 70;
  Best Local Similarity   42.9%;  Pred. No. 4.1e+04;
  Matches    9;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |:::   :|| ||| | |
Db         15 ATACTTTATTTAATCCCAAAG 35


RESULT 3
BQ613481/c
LOCUS       BQ613481                  72 bp    mRNA    linear   EST 26-JUN-2002
DEFINITION  rd07h06.y1 Meloidogyne incognita egg SL1 TOPO v1 Meloidogyne
            incognita cDNA 5', mRNA sequence.
ACCESSION   BQ613481
VERSION     BQ613481.1  GI:21603157
KEYWORDS    EST.
SOURCE      Meloidogyne incognita (southern root-knot nematode)
  ORGANISM  Meloidogyne incognita
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. Meloidogyne incognita eggs
            were provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
            Putative full length read
            The vector to vector length is 73
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Meloidogyne incognita"
                     /mol_type="mRNA"
                     /db_xref="taxon:6306"
                     /dev_stage="egg"
                     /lab_host="DH10B (Invitrogen)"
                     /clone_lib="Meloidogyne incognita egg SL1 TOPO v1"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy and Dr. James McCarter at Washington University,
                     St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR
                     products of size >400 nucleotides containing SL1 on the 5'
                     end and oligo(dT) on the 3' end were non-directionally
                     cloned into pCRII-TOPO(Invitrogen) following the Topo TA
                     cloning protocol. Meloidogyne incognita eggs were provided
                     by Andrew Kloek of Divergence Inc., St. Louis, MO."
ORIGIN

  Query Match             44.1%;  Score 12.8;  DB 13;  Length 72;
  Best Local Similarity   40.9%;  Pred. No. 4.1e+04;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANGNG 27
              :  :::   :||||| | | |
Db         50 TCTTTTTTCTAAGCCCCACGGG 29


RESULT 4
AQ025263/c
LOCUS       AQ025263                  76 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(3)3081 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from 5' end of P element, genomic survey
            sequence.
ACCESSION   AQ025263
VERSION     AQ025263.1  GI:3265615
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
   PUBMED   10716700
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 69 in the 76 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
```

V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."

ORIGIN

```
  Query Match             44.1%;  Score 12.8;  DB 28;  Length 76;
  Best Local Similarity   42.9%;  Pred. No. 4.2e+04;
  Matches    9;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |:::   :|| ||| | |
Db         62 ATACTTTATTTAATCCCAAAG 42
```

RESULT 5
AA700959/c

LOCUS	AA700959 52 bp mRNA linear EST 19-DEC-1997
DEFINITION	zf87d10.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone IMAGE:383923 3' similar to TR:P79324 P79324 RIBOSOMAL PROTEIN L15 ;, mRNA sequence.
ACCESSION	AA700959
VERSION	AA700959.1 GI:2704124
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 52)
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE	WashU-NCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 1.

FEATURES	Location/Qualifiers
source	1. .52 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="GDB:1292180" /db_xref="taxon:9606" /clone="IMAGE:383923" /lab_host="DH10B (ampicillin resistant)" /clone_lib="Soares_pineal_gland_N3HPG" /note="Organ: pineal gland; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. "

ORIGIN

```
  Query Match             42.8%;  Score 12.4;  DB 9;  Length 52;
  Best Local Similarity   52.9%;  Pred. No. 6.4e+04;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: | ::  |:|||||
Db         31 GATCCATTCCGTAAGCC 15
```

RESULT 6
AA468615/c

LOCUS	AA468615 70 bp mRNA linear EST 15-AUG-1997
DEFINITION	ne08b04.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:880591 3' similar to TR:G178246 G178246 MATURE ALPHA-GALACTOSIDASE A, COMPLETE CDS PRECURSOR ;, mRNA sequence.
ACCESSION	AA468615
VERSION	AA468615.1 GI:2195149
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 70)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert Length: 1285 Std Error: 0.00 Seq primer: -41m13 fwd. ET from Amersham High quality sequence stop: 1.

FEATURES	Location/Qualifiers
source	1. .70 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:880591" /sex="pooled" /tissue_type="colon" /lab_host="DH10B" /clone_lib="NCI_CGAP_Co3" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. "

ORIGIN

```
  Query Match             42.8%;  Score 12.4;  DB 9;  Length 70;
  Best Local Similarity   52.9%;  Pred. No. 6.6e+04;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: |:::  | |||||
Db         64 GATCCTTTGGGCAAGCC 48
```

RESULT 7
AZ503560

LOCUS	AZ503560 48 bp DNA linear GSS 05-OCT-2000
DEFINITION	1M0343E21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0343E21 F, genomic survey sequence.
ACCESSION	AZ503560
VERSION	AZ503560.1 GI:10684876
KEYWORDS	GSS.

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0343  row: E  column: 21
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0343E21"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 28;  Length 48;
  Best Local Similarity   43.5%;  Pred. No. 8e+04;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: |: :  |:||| |  | | |
Db         13 ATGCTCTGGGTAAGGCACAAGAG 35


RESULT 8
BZ770420
LOCUS       BZ770420                 56 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  SALK_143355.56.00.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_143355.56.00.x, genomic
            survey sequence.
ACCESSION   BZ770420
VERSION     BZ770420.1  GI:28944104
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At5g40030.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_143355.56.00.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 28;  Length 56;
  Best Local Similarity   45.5%;  Pred. No. 8.1e+04;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |: :   :|||||  | |
Db          3 GATACTATTGTTAAGCCTAACG 24


RESULT 9
BX744082
LOCUS       BX744082                 66 bp    mRNA    linear   EST 18-NOV-2003
DEFINITION  BX744082 XGC-tadpole Silurana tropicalis cDNA clone TTpA072h19 3',
            mRNA sequence.
ACCESSION   BX744082
VERSION     BX744082.1  GI:38416822
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TTpA072h19.q1kaT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
```

```
            constructed by Nigel Garrett.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from tadpole
            embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli DH10B.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Silurana tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TTpA072h19"
                     /dev_stage="tadpole (stage 35-40)"
                     /lab_host="E. coli DH10B"
                     /clone_lib="XGC-tadpole"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from tadpole
                     embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
                     with EcoRI at the 5' end and NotI at the 3' end"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 13;  Length 66;
  Best Local Similarity   43.5%;  Pred. No. 8.3e+04;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: |:::   :|| ||| | | |
Db         20 ATGCCTTATTTAATCCCCATGTG 42


RESULT 10
CD946435/c
LOCUS       CD946435                 67 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  REN_47 GeneTag1 Zea mays cDNA, mRNA sequence.
ACCESSION   CD946435
VERSION     CD946435.1  GI:32794199
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="mixture"
                     /db_xref="taxon:4577"
                     /clone_lib="GeneTag1"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 14;  Length 67;
  Best Local Similarity   50.0%;  Pred. No. 8.4e+04;
  Matches   11;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |: :  | | ||||  | |
Db         25 GATACTCTGGGGATGCCCTAAG 4


RESULT 11
AZ453746/c
LOCUS       AZ453746                 75 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0255A23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0255A23 F, genomic survey sequence.
ACCESSION   AZ453746
VERSION     AZ453746.1  GI:10611850
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0255  row: A  column: 23
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 75.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0255A23"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 28;  Length 75;
  Best Local Similarity   43.5%;  Pred. No. 8.5e+04;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |:  : :  |:||| || | | |
Db         24 ATAATGTTTGTAAGTCCAATGGG 2


RESULT 12
```

```
AA975071/c
LOCUS       AA975071                 40 bp    mRNA    linear   EST 26-AUG-1998
DEFINITION  on03d07.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1555597 3'
            similar to TR:P70566 P70566 N-TROPOMODULIN. ;, mRNA sequence.
ACCESSION   AA975071
VERSION     AA975071.1  GI:3150863
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 2096   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1555597"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid3"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo. "
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 9;  Length 40;
  Best Local Similarity   40.0%;  Pred. No. 1.2e+05;
  Matches    8;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              : |:::  |:||| ||   |
Db         35 TCCTTTCCGTAAGACCTTGG 16


RESULT 13
BE970036/c
LOCUS       BE970036                 49 bp    mRNA    linear   EST 04-OCT-2000
DEFINITION  601680150F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950172 5',
            mRNA sequence.
ACCESSION   BE970036
VERSION     BE970036.1  GI:10582969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM816  row: d  column: 13
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3950172"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_78"
                     /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); 5' and 3' adaptors were used in cloning
                     as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
                     and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.2
                     kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA)."
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 49;
  Best Local Similarity   45.0%;  Pred. No. 1.3e+05;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              :  :::  | ||||||  | |
Db         48 TTTTTTATGCAAGCCCCAGG 29


RESULT 14
CC516004
LOCUS       CC516004                 51 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_361F9.T7 CHORI-240 Bos taurus genomic clone CH240_361F9,
            genomic survey sequence.
ACCESSION   CC516004
VERSION     CC516004.1  GI:31834292
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_361F9.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
```

```
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 361  row: F  column: 9
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_361F9"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 51;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+05;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||:  :::  |: |||||
Db         31 GATGATTTCAGTGAGCCC 48


RESULT 15
CG563472/c
LOCUS       CG563472                  63 bp    DNA     linear   GSS 01-OCT-2003
DEFINITION  OST186777 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST186777, genomic survey sequence.
ACCESSION   CG563472
VERSION     CG563472.1  GI:37350059
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST186777"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 63;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              | : |: :  |:||||||
Db         45 GTTTCTCTGTGTAAGCCC 28


RESULT 16
AA733449/c
LOCUS       AA733449                  65 bp    mRNA    linear   EST 07-JAN-1998
DEFINITION  vt73h08.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1176831 5' similar to gb:X06617 40S RIBOSOMAL PROTEIN
            S11 (HUMAN);, mRNA sequence.
ACCESSION   AA733449
VERSION     AA733449.1  GI:2755116
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:634679
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1176831"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG], digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead.  "
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 9;  Length 65;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+05;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy          4 GAUNCUUUNNGUAAGCCC 21
              | : |:::  |:|||| |
Db         44 GCTGCTTTTGGTAAGCAC 27
```

RESULT 17
AI767928
LOCUS AI767928 70 bp mRNA linear EST 21-DEC-1999
DEFINITION wi99c01.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2401440 3' similar to SW:RT14_HUMAN P78537 RT14 PROTEIN ;, mRNA sequence.
ACCESSION AI767928
VERSION AI767928.1 GI:5234426
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 574 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2401440"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 40.7%; Score 11.8; DB 9; Length 70;
Best Local Similarity 47.4%; Pred. No. 1.4e+05;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCCCNA 23
              |: |:::  | |||||  |
Db         46 ATTCTTTAAGCAAGCCAGA 64
```

RESULT 18
BH216023/c
LOCUS BH216023 70 bp DNA linear GSS 08-NOV-2001
DEFINITION 1006039G04.2EL_y1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.
ACCESSION BH216023
VERSION BH216023.1 GI:16806681
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 70)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006039 row: 43
Class: transposon-tagged.

FEATURES Location/Qualifiers
source 1. .70
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 40.7%; Score 11.8; DB 28; Length 70;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||: |:::  | | ||||
Db         70 GATCCTTTTAGGAGGCCC 53
```

RESULT 19
BX534001
LOCUS BX534001 72 bp DNA linear GSS 03-JUN-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-505G09-019706, genomic survey sequence.
ACCESSION BX534001
VERSION BX534001.1 GI:31411131
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA

```
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 72)
  AUTHORS   Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2003) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone T3H13. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-505G09-019706"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 72;
  Best Local Similarity   47.4%;  Pred. No. 1.4e+05;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNA 23
              |:  ::   |:|||||  |
Db         22 ATCGTTGCTGTAAGCCCAA 40


RESULT 20
CG574740
LOCUS       CG574740                  74 bp    DNA     linear   GSS 02-OCT-2003
DEFINITION  OST207881 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST207881, genomic survey sequence.
ACCESSION   CG574740
VERSION     CG574740.1  GI:37365077
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST207881"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 74;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+05;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              | : |::   |:||||||
Db          2 GCTACTTCGTGTAAGCCC 19


RESULT 21
CNS01561/c
LOCUS       CNS01561                  75 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN13J24 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL105043
VERSION     AL105043.1  GI:5617057
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN13J24"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 75;
  Best Local Similarity   45.8%;  Pred. No. 1.4e+05;
  Matches   11;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  | :   |:|  ||| | | |
Db         30 GANGCNTNCGGTATCCCCCAGGGG 7
```

```
RESULT 22
AV962932
LOCUS       AV962932                 76 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  AV962932 Nori Satoh unpublished cDNA library, cleavage stage embryo
            Ciona intestinalis cDNA clone cicl22b09 5', mRNA sequence.
ACCESSION   AV962932
VERSION     AV962932.1  GI:19451231
KEYWORDS    EST.
SOURCE      Ciona intestinalis
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="cicl22b09"
                     /tissue_type="whole animal"
                     /dev_stage="cleavage stage embryo"
                     /clone_lib="Nori Satoh unpublished cDNA library, cleavage
                     stage embryo"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 9;  Length 76;
  Best Local Similarity   38.1%;  Pred. No. 1.4e+05;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |:::  |: | | | | |
Db         36 ATGCTTTGCGTNAACTCAAAG 56


RESULT 23
BZ289518/c
LOCUS       BZ289518                 76 bp    DNA     linear   GSS 24-OCT-2002
DEFINITION  SALK_022917.29.15.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_022917.29.15.x, genomic
            survey sequence.
ACCESSION   BZ289518
VERSION     BZ289518.1  GI:24331254
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_022917.29.15.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 28;  Length 76;
  Best Local Similarity   47.4%;  Pred. No. 1.4e+05;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNA 23
              |: |::   |:||| || |
Db         76 ATTCTTGATGTAAGGCCCA 58


RESULT 24
BQ100875
LOCUS       BQ100875                 77 bp    mRNA    linear   EST 29-APR-2002
DEFINITION  ij25a03.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA clone IMAGE:6135461 3', mRNA sequence.
ACCESSION   BQ100875
VERSION     BQ100875.1  GI:20133859
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ij25a03.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center This clone is
            available royalty-free through LLNL; please contact the IMAGE
            consortium (info@image.llnl.gov) for further information
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6135461"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
```

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

ORIGIN

Query Match 40.7%; Score 11.8; DB 13; Length 77;
Best Local Similarity 45.0%; Pred. No. 1.4e+05;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy          6 UNCUUUNNGUAAGCCCNANG 25
              :  :::  | ||||||  | |
Db          8 TTTTTTCTGCAAGCCCCAGG 27
```

RESULT 25
BH904967
LOCUS BH904967 78 bp DNA linear GSS 04-SEP-2002
DEFINITION SALK_105401.55.70.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_105401.55.70.x, genomic survey sequence.
ACCESSION BH904967
VERSION BH904967.1 GI:22717698
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 78)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.
Class: TDNA tagged.

FEATURES Location/Qualifiers
source 1. .78
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_105401.55.70.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 40.7%; Score 11.8; DB 28; Length 78;
Best Local Similarity 42.1%; Pred. No. 1.4e+05;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCCCNA 23
              |:  :::  |:|| ||| |
Db         36 ATTATTTTCGTAACCCCAA 54
```

RESULT 26
CG573962/c
LOCUS CG573962 79 bp DNA linear GSS 02-OCT-2003
DEFINITION OST206260 Mus musculus 129Sv/Ev Mus musculus genomic clone OST206260, genomic survey sequence.
ACCESSION CG573962
VERSION CG573962.1 GI:37364299
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 79)
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J., Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,C. and Sands,A.T.
TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES Location/Qualifiers
source 1. .79
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129SV/EV"
/db_xref="taxon:10090"
/clone="OST206260"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 40.7%; Score 11.8; DB 29; Length 79;
Best Local Similarity 38.1%; Pred. No. 1.4e+05;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |:::  |: ||| |   |
Db         32 ATGCTTTGTGTCAGCNCCTCG 12
```

RESULT 27
AL951473
LOCUS AL951473 79 bp DNA linear GSS 24-OCT-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-336E04-016157, genomic survey sequence.
ACCESSION AL951473
VERSION AL951473.1 GI:24408095
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1

```
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 79)
  AUTHORS   Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone T25N20.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-336E04-016157"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 79;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+05;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||: |:::  |:||  ||
Db          8 GATTCTTTACGTAAAACC 25


RESULT 28
U44334
LOCUS       U44334                    49 bp    mRNA    linear   EST 03-APR-1996
DEFINITION  ENU44334 Aspergillus nidulans cleistothecium Emericella nidulans
            cDNA clone SE0762, mRNA sequence.
ACCESSION   U44334
VERSION     U44334.1  GI:1244997
KEYWORDS    EST.
SOURCE      Emericella nidulans (anamorph: Aspergillus nidulans)
  ORGANISM  Emericella nidulans
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; Emericella.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Lee,D., Lee,S., Hwang,H., Kim,J. and Chae,K.
  TITLE     Quantitative analysis of gene expression in sexual structures of
            Aspergillus nidulans by sequencing of 3'-directed cDNA clones
  JOURNAL   FEMS Microbiol. Lett. 138 (1), 71-76 (1996)
  MEDLINE   96236220
   PUBMED   8674973
COMMENT     Contact: Keon-Sang Chae
            Chonbuk National University
            Chonju, 561-756, S. Korea
            Tel: +82-652-70-3340
            Fax: +82-652-70-3345
            Email: chaeks@chonbuknms.chonbuk.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Emericella nidulans"
                     /mol_type="mRNA"
                     /strain="FGSC4"
                     /db_xref="taxon:162425"
                     /clone="SE0762"
                     /tissue_type="cleistothecium"
                     /cell_type="Hull cell"
                     /dev_stage="sexual"
                     /clone_lib="Aspergillus nidulans cleistothecium"
                     /note="3'-directed cDNA clones; single-pass sequencing"
ORIGIN

  Query Match             40.0%;  Score 11.6;  DB 14;  Length 49;
  Best Local Similarity   37.5%;  Pred. No. 1.6e+05;
  Matches    9;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::   :||  || |   |
Db         21 GATTCTTTCATTAACTCCCAACGG 44


RESULT 29
AI584456/c
LOCUS       AI584456                  58 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fb93h12.x1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3719495 3' similar to SW:TRF1_SALSA P80426 SEROTRANSFERRIN I
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI584456
VERSION     AI584456.1  GI:4570353
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 1
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
```

/clone="IMAGE:3719495"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XL1-blue MRF"
/clone_lib="Zebrafish WashU MPIMG EST"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik,Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

ORIGIN

Query Match 40.0%; Score 11.6; DB 9; Length 58;
Best Local Similarity 37.5%; Pred. No. 1.7e+05;
Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              | : |:::   :|||| |   | |
Db         25 GTTGCTTTTATTAAGCACACTGTG 2
```

RESULT 30
CG519587/c
LOCUS CG519587 65 bp DNA linear GSS 01-OCT-2003
DEFINITION OST83436 Mus musculus 129Sv/Ev Mus musculus genomic clone OST83436, genomic survey sequence.
ACCESSION CG519587
VERSION CG519587.1 GI:37306159
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 65)
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J., Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,C. and Sands,A.T.
TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
FEATURES Location/Qualifiers
source 1. .65
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129SV/EV"
/db_xref="taxon:10090"
/clone="OST83436"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 40.0%; Score 11.6; DB 29; Length 65;
Best Local Similarity 45.8%; Pred. No. 1.7e+05;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |: :  | |||||    | |
Db         27 GATGCTGTAAGAAAGCCAGCTGAG 4
```

RESULT 31
AI814489
LOCUS AI814489 70 bp mRNA linear EST 24-AUG-1999
DEFINITION wj73g11.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2408516 3' similar to gb:X59268 TRANSCRIPTION INITIATION FACTOR IIB (HUMAN);, mRNA sequence.
ACCESSION AI814489
VERSION AI814489.1 GI:5425704
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2408516"
/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Lu19"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went througb one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 40.0%; Score 11.6; DB 9; Length 70;
Best Local Similarity 45.8%; Pred. No. 1.7e+05;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
```

```
                ||  |: :  |:|||  | | | |
Db          24 GAGTCTCTTGGTAAGGGCGATGAG 47


RESULT 32
BX001193
LOCUS       BX001193                73 bp    DNA     linear   GSS 04-DEC-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-104G08-018339,
            genomic survey sequence.
ACCESSION   BX001193
VERSION     BX001193.1  GI:26186153
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 73)
  AUTHORS   Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone T22B15.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-104G08-018339"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             40.0%;  Score 11.6;  DB 29;  Length 73;
  Best Local Similarity   41.7%;  Pred. No. 1.7e+05;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||:  :::    || ||| | | |
Db         46 GATATTTTATAGAAACCCTATGGG 69


RESULT 33
CG549254
LOCUS       CG549254                77 bp    DNA     linear   GSS 01-OCT-2003
DEFINITION  OST152329 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST152329, genomic survey sequence.
ACCESSION   CG549254
VERSION     CG549254.1  GI:37335841
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST152329"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             40.0%;  Score 11.6;  DB 29;  Length 77;
  Best Local Similarity   37.5%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::   :||||     | |
Db          8 GATTCTTTTTCTAAGCAAGCTGGG 31


RESULT 34
AL759596/c
LOCUS       AL759596                77 bp    DNA     linear   GSS 18-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-188G10-014624,
            genomic survey sequence.
ACCESSION   AL759596
VERSION     AL759596.1  GI:21497944
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 77)
```

```
  AUTHORS   Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone MRI1. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-188G10-014624"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             40.0%;  Score 11.6;  DB 29;  Length 77;
  Best Local Similarity   41.7%;  Pred. No. 1.7e+05;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |: :  |:| |||  |   |
Db         74 GATTCTATAGGTAGGCCTAATTGG 51


RESULT 35
CF115782/c
LOCUS       CF115782                 79 bp    mRNA    linear   EST 23-JUL-2003
DEFINITION  SSFP015D17-TplX5.z1 fp 73-75d fetal sheep skin library Ovis aries
            cDNA, mRNA sequence.
ACCESSION   CF115782
VERSION     CF115782.1  GI:33177048
KEYWORDS    EST.
SOURCE      Ovis aries (sheep)
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   DeSilva,U., Franklin,I.R., Maddox,J.F., van Hest,B. and
            Adelson,D.L.
  TITLE     Gene Expression in Sheep Skin and Wool (Hair)
  JOURNAL   Cytogenet. Genome Res. (2003) In press
COMMENT     Contact: David L. Adelson
            Animal Breeding and Genetics
            Texas A&M University
            Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
            USA
            Tel: 9798452616
            Fax: 9798456970
            Email: david.adelson@tamu.edu.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Ovis aries"
                     /mol_type="mRNA"
                     /strain="Merino"
                     /db_xref="taxon:9940"
                     /sex="male"
                     /dev_stage="73-75d pc fetus"
                     /lab_host="BM25.8"
                     /clone_lib="fp 73-75d fetal sheep skin library"
                     /note="Organ: midflank skin; Vector: pTriplEX; Site_1:
                     EcoRI; Site_2: XbaI; Arrayed library filters screened by
                     hybridization with labeled dinucleotide repeats. first
                     strand synthesis XbaI-(dT)15 primed, EcoRI adaptors were
                     ligated to cDNA before cloning. EcoRI- NotI- SalI adaptor
                     sequence;5'-AATTCGCGGCCGCGTCGAC    GCGCCGGCGCAGCTG-PO4-5'
                     5' sequencing primer 5' TCCGAGATCTGGACGAGC 3' 3'
                     sequencing primer 5' TAATACGACTCACTATAGGG 3' submitted
                     sequences in excess of 50 bp after vector and quality
                     clipping(phred 13) submitted to GenBank."
ORIGIN

  Query Match             40.0%;  Score 11.6;  DB 14;  Length 79;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+05;
  Matches   12;  Conservative    2;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||    ::  | ||||||  | | |
Db         76 GAAAAATTGGGGAAGCCCAAAGGG 53


RESULT 36
AZ336769/c
LOCUS       AZ336769                 79 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0067L12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0067L12 F, genomic survey sequence.
ACCESSION   AZ336769
VERSION     AZ336769.1  GI:10406402
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0067  row: L  column: 12
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 79.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0067L12"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
```

polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 40.0%; Score 11.6; DB 28; Length 79;
Best Local Similarity 41.7%; Pred. No. 1.7e+05;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              | : |:::    |||||  | | |
Db         79 GCTACTTTGCTAAAGCCATAGGGG 56
```

RESULT 37
HSMC42B09/c
LOCUS HSMC42B09 32 bp DNA linear GSS 29-MAY-1997
DEFINITION H.sapiens DNA for trapped exon (ID HMC42B09), genomic survey sequence.
ACCESSION X88068
VERSION X88068.1 GI:1437990
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 32)
AUTHORS Chen,H., Chrast,R., Rossier,C., Morris,M.A., Lalioti,M.D. and Antonarakis,S.E.
TITLE Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping
JOURNAL Genome Res. 6 (8), 747-760 (1996)
MEDLINE 97011340
PUBMED 8858350
REFERENCE 2 (bases 1 to 32)
AUTHORS Chen,H.M., Rossier,C., Chrast,R. and Antonarakis,S.E.
TITLE Cloning of trapped exons from human chromosome 21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 32)
AUTHORS Antonarakis,S.E.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of Medical Genetics, University and Cantonal Hospital of Geneva, CMU, 1 rue Michel-Servet, 1211 Geneva, SWITZERLAND
FEATURES Location/Qualifiers
source 1. .32
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
exon 1. .32
/note="trapped exon"
ORIGIN

Query Match 39.3%; Score 11.4; DB 29; Length 32;
Best Local Similarity 47.1%; Pred. No. 1.9e+05;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: |:::    |||||
Db         31 GATACTTTCANCAAGCC 15
```

RESULT 38
BX001854
LOCUS BX001854 34 bp DNA linear GSS 04-DEC-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-359D06-016879, genomic survey sequence.
ACCESSION BX001854
VERSION BX001854.1 GI:26186814
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
JOURNAL Unpublished
REFERENCE 2
AUTHORS Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 34)
AUTHORS Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g13370. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES Location/Qualifiers
source 1. .34
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-359D06-016879"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thalaina nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"
ORIGIN

Query Match 39.3%; Score 11.4; DB 29; Length 34;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCC 20
              |: |: :  |:|||||
Db         12 ATCCTCTGTGTAAGCC 27
```

RESULT 39
TA253H01Q/c
LOCUS TA253H01Q 40 bp DNA linear GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 253h01, reverse sequence, genomic survey sequence.
ACCESSION AL483109
VERSION AL483109.1 GI:11849577
KEYWORDS GSS.
SOURCE Trypanosoma brucei

```
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="Trypanosoma brucei"
                      /mol_type="genomic DNA"
                      /strain="TREU927"
                      /db_xref="taxon:5691"
                      /clone="253h01"
ORIGIN

  Query Match             39.3%;  Score 11.4;  DB 29;  Length 40;
  Best Local Similarity   55.6%;  Pred. No. 2e+05;
  Matches   10;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         10 UUNNGUAAGCCCNANGNG 27
              ::  | ||||||  | | |
Db         31 TTATGCAAGCCCCATGCG 14


RESULT 40
AI887082
LOCUS       AI887082                 46 bp    mRNA    linear   EST 07-MAR-2000
DEFINITION  wl96e09.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2432776 3'
            similar to TR:O60869 O60869 EDF-1 PROTEIN. ;, mRNA sequence.
ACCESSION   AI887082
VERSION     AI887082.1  GI:5592246
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 749   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2432776"
                      /tissue_type="anaplastic oligodendroglioma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Brn25"
                      /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      Library is normalized, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             39.3%;  Score 11.4;  DB 9;  Length 46;
  Best Local Similarity   47.1%;  Pred. No. 2e+05;
  Matches    8;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 UUUNNGUAAGCCCNANG 25
              :::  |:||||||   |
Db         25 TTTGGGTAAGCCCCTTG 41


Search completed: April 18, 2004, 09:58:23
Job time : 1591.67 secs
```

OM nucleic - nucleic search, using sw model

Run on: April 18, 2004, 05:05:34 ; Search time 179.667 Seconds
(without alignments)
685.702 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3399856

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	29	3	AAA70829	Aaa70829 Molecular
2	29	100.0	29	3	AAA70830	Aaa70830 Molecular
3	29	100.0	42	3	AAA71121	Aaa71121 Molecular
4	29	100.0	42	3	AAA71128	Aaa71128 Molecular
5	29	100.0	42	3	AAA71120	Aaa71120 Molecular
6	29	100.0	42	3	AAA71116	Aaa71116 Molecular
7	29	100.0	42	3	AAA71115	Aaa71115 Molecular
8	29	100.0	42	3	AAA71129	Aaa71129 Molecular
9	28	96.6	45	3	AAA70826	Aaa70826 Molecular
10	28	96.6	45	3	AAA70825	Aaa70825 Molecular
11	28	96.6	46	3	AAA71089	Aaa71089 Molecular
12	28	96.6	46	3	AAA71106	Aaa71106 Molecular
13	28	96.6	46	3	AAA71107	Aaa71107 Molecular
14	28	96.6	46	3	AAA71088	Aaa71088 Molecular
15	28	96.6	46	3	AAA71105	Aaa71105 Molecular
16	28	96.6	46	3	AAA71090	Aaa71090 Molecular
17	24.8	85.5	42	3	AAA71113	Aaa71113 Molecular
18	24.8	85.5	42	3	AAA71118	Aaa71118 Molecular
19	24.8	85.5	42	3	AAA71126	Aaa71126 Molecular
20	23.8	82.1	46	3	AAA71085	Aaa71085 Molecular
21	23.8	82.1	46	3	AAA71103	Aaa71103 Molecular
22	23.2	80.0	29	3	AAA70828	Aaa70828 Molecular
23	23.2	80.0	42	3	AAA71123	Aaa71123 Molecular
24	23.2	80.0	42	3	AAA71131	Aaa71131 Molecular
25	22.2	76.6	45	3	AAA70824	Aaa70824 Molecular
26	22.2	76.6	46	3	AAA71087	Aaa71087 Molecular
27	22.2	76.6	46	3	AAA71096	Aaa71096 Molecular
28	22.2	76.6	46	3	AAA71099	Aaa71099 Molecular
29	22.2	76.6	46	3	AAA71100	Aaa71100 Molecular
30	22.2	76.6	46	3	AAA71104	Aaa71104 Molecular
31	21.2	73.1	42	3	AAA71114	Aaa71114 Molecular
32	21.2	73.1	42	3	AAA71119	Aaa71119 Molecular
33	21.2	73.1	42	3	AAA71127	Aaa71127 Molecular
34	21.2	73.1	46	3	AAA71094	Aaa71094 Molecular
35	21.2	73.1	46	3	AAA71110	Aaa71110 Molecular
36	20	69.0	46	3	AAA71098	Aaa71098 Molecular
37	20	69.0	46	3	AAA71102	Aaa71102 Molecular
38	20	69.0	46	3	AAA71084	Aaa71084 Molecular
39	19.6	67.6	42	3	AAA71124	Aaa71124 Molecular
40	19.6	67.6	42	3	AAA71132	Aaa71132 Molecular
41	18.6	64.1	46	3	AAA71111	Aaa71111 Molecular
42	18.6	64.1	46	3	AAA71095	Aaa71095 Molecular
43	18.6	64.1	46	3	AAA71109	Aaa71109 Molecular
44	18.6	64.1	46	3	AAA71093	Aaa71093 Molecular
45	18.4	63.4	42	3	AAA71130	Aaa71130 Molecular
46	18.4	63.4	42	3	AAA71117	Aaa71117 Molecular
47	18.4	63.4	42	3	AAA71122	Aaa71122 Molecular
48	18	62.1	29	3	AAA70827	Aaa70827 Molecular
49	18	62.1	44	3	ABK87476	Abk87476 Interleuk
50	18	62.1	44	3	AAA71112	Aaa71112 Molecular
51	18	62.1	44	3	AAA71125	Aaa71125 Molecular
52	18	62.1	44	3	AAA71133	Aaa71133 Molecular
c 53	16.4	56.6	31	6	ABK21678	Abk21678 Human ERG
c 54	16.2	55.9	48	3	AAZ47019	Aaz47019 Primer JC
55	16.2	55.9	79	3	AAC25823	Aac25823 Human sec
c 56	15.6	53.8	51	4	AAL29619	Aal29619 Human SNP
c 57	15.4	53.1	41	6	ABZ50144	Abz50144 Human NDU
c 58	15.4	53.1	41	6	ABZ44134	Abz44134 Human NDU
59	15.4	53.1	53	7	ABX54825	Abx54825 Bovine ES
c 60	15.2	52.4	33	2	AAV32805	Aav32805 Forward p
c 61	15.2	52.4	60	6	ABN45977	Abn45977 Human spl
c 62	15.2	52.4	75	2	AAQ62654	Aaq62654 Tobacco-m
63	15	51.7	23	2	AAT01534	Aat01534 Human her
64	15	51.7	23	2	AAT03719	Aat03719 Human her
65	15	51.7	30	6	ABX69578	Abx69578 Novel Hel
c 66	15	51.7	33	4	AAH19846	Aah19846 Sense pri
c 67	15	51.7	33	7	ABT14538	Abt14538 Universal
68	15	51.7	39	7	ABX99212	Abx99212 Human CAN
c 69	15	51.7	41	6	ABL54589	Abl54589 Human pro
c 70	15	51.7	60	6	ABN39439	Abn39439 Human spl
71	15	51.7	65	6	ABN31832	Abn31832 Rat splic
c 72	14.8	51.0	29	2	AAT18757	Aat18757 Primer #1
c 73	14.8	51.0	29	2	AAT63158	Aat63158 Primer 1,
c 74	14.8	51.0	29	2	AAT93525	Aat93525 Locus spe
c 75	14.8	51.0	29	2	AAV06206	Aav06206 Primer us
c 76	14.8	51.0	29	2	AAV36871	Aav36871 Nucleotid
c 77	14.8	51.0	29	3	AAA47262	Aaa47262 Primer 1
c 78	14.8	51.0	32	2	AAV06226	Aav06226 Primer us
c 79	14.8	51.0	32	3	AAA47264	Aaa47264 Primer 3
c 80	14.8	51.0	41	6	ABL54588	Abl54588 Human pro
c 81	14.8	51.0	46	6	ABN72001	Abn72001 Streptoco
82	14.8	51.0	75	7	ABZ79961	Abz79961 Potexviru
83	14.8	51.0	75	7	ABZ79980	Abz79980 Potexviru
c 84	14.8	51.0	80	2	AAT25584	Aat25584 Human gen
85	14.8	51.0	80	2	AAT18556	Aat18556 Human cho
86	14.6	50.3	25	3	AAC96431	Aac96431 HLA DQA1
87	14.6	50.3	25	6	AAD37634	Aad37634 Trichoder
88	14.6	50.3	25	7	ACD29574	Acd29574 F22844 ly
89	14.6	50.3	50	4	AAL29124	Aal29124 Human SNP
c 90	14.6	50.3	53	2	AAT47169	Aat47169 Primer JC
91	14.6	50.3	60	6	ABN45366	Abn45366 Human spl
92	14.6	50.3	60	6	ABN35687	Abn35687 Human spl
93	14.6	50.3	60	7	ABZ82291	Abz82291 Erythropo
c 94	14.6	50.3	73	3	AAC21624	Aac21624 Human sec
c 95	14.6	50.3	75	4	AAI25979	Aai25979 Probe #15
c 96	14.6	50.3	75	4	ABA72971	Aba72971 Human foe

c 97	14.6	50.3	75	4	AAI53395	Aai53395 Probe #22
c 98	14.6	50.3	75	4	ABA38521	Aba38521 Probe #16
c 99	14.6	50.3	75	4	AAK47563	Aak47563 Human bon
c 100	14.6	50.3	75	4	AAK21402	Aak21402 Human bra
c 101	14.6	50.3	75	4	ABS47298	Abs47298 Human liv
c 102	14.6	50.3	75	6	ABS21657	Abs21657 Human gen
103	14.4	49.7	25	3	AAC96808	Aac96808 HLA HLA-C
c 104	14.4	49.7	31	6	ABK21667	Abk21667 Human ERG
c 105	14.4	49.7	31	7	ACD58557	Acd58557 HCV DNAzy
106	14.4	49.7	51	4	AAH39512	Aah39512 Human SNP
c 107	14.4	49.7	58	2	AAT24569	Aat24569 Human gen
108	14.4	49.7	65	6	ABN57807	Abn57807 Mouse spl
c 109	14.4	49.7	65	6	ABN31176	Abn31176 Rat splic
c 110	14.2	49.0	25	6	ABS75578	Abs75578 Human PAP
c 111	14.2	49.0	25	6	ABS75577	Abs75577 Human PAP
c 112	14.2	49.0	25	6	ABS75573	Abs75573 Human PAP
c 113	14.2	49.0	25	6	ABS75572	Abs75572 Human PAP
c 114	14.2	49.0	25	6	ABS75576	Abs75576 Human PAP
c 115	14.2	49.0	25	6	ABS75574	Abs75574 Human PAP
c 116	14.2	49.0	25	6	ABS75575	Abs75575 Human PAP
117	14.2	49.0	25	8	ACI76453	Aci76453 Human mic
c 118	14.2	49.0	30	6	ABX68475	Abx68475 Novel Hel
c 119	14.2	49.0	31	6	ABK60044	Abk60044 Human CLC
c 120	14.2	49.0	31	7	ACD62445	Acd62445 HCV minus
c 121	14.2	49.0	42	2	AAX34817	Aax34817 Human ZSI
c 122	14.2	49.0	42	3	AAA71113	Aaa71113 Molecular
c 123	14.2	49.0	42	3	AAA71118	Aaa71118 Molecular
c 124	14.2	49.0	42	3	AAA71126	Aaa71126 Molecular
125	14.2	49.0	50	4	AAL28930	Aal28930 Human SNP
c 126	14.2	49.0	50	6	ABZ02089	Abz02089 Human leu
c 127	14.2	49.0	51	4	AAL31644	Aal31644 Human SNP
c 128	14.2	49.0	51	4	AAL27753	Aal27753 Human SNP
c 129	14.2	49.0	51	4	AAL26990	Aal26990 Human SNP
130	14.2	49.0	55	2	AAQ37152	Aaq37152 Probe to
c 131	14.2	49.0	60	6	ABN41192	Abn41192 Human spl
c 132	14.2	49.0	60	6	ABN44060	Abn44060 Human spl
133	14.2	49.0	60	6	ABN41092	Abn41092 Human spl
c 134	14.2	49.0	60	9	ADE87528	Ade87528 Bovine la
135	14.2	49.0	65	6	ABZ26621	Abz26621 Candida e
136	14.2	49.0	69	2	AAV08003	Aav08003 Probe IL-
137	14	48.3	25	3	AAC96192	Aac96192 16s rRNA
138	14	48.3	41	6	AAD33378	Aad33378 pBADmyc/h
139	14	48.3	41	6	ABS56697	Abs56697 Human rib
140	14	48.3	41	6	ABS56696	Abs56696 Human rib
141	14	48.3	41	10	ADE48045	Ade48045 T4 RNA li
142	14	48.3	47	3	AAZ66750	Aaz66750 Human map
c 143	14	48.3	51	4	AAL33702	Aal33702 Human SNP
144	14	48.3	54	6	ABX05557	Abx05557 Secondary
c 145	13.8	47.6	25	5	AAS22104	Aas22104 Human COL
c 146	13.8	47.6	27	5	AAS17481	Aas17481 Tobacco m
c 147	13.8	47.6	27	6	ABK87703	Abk87703 Kinased 5
c 148	13.8	47.6	27	9	AAD62031	Aad62031 PCR prime
c 149	13.8	47.6	28	3	AAA06812	Aaa06812 PCR prime
150	13.8	47.6	30	6	ABX67821	Abx67821 Novel Hel
c 151	13.8	47.6	31	7	ACD65561	Acd65561 HCV minus
c 152	13.8	47.6	33	3	AAA47266	Aaa47266 Primer 5
c 153	13.8	47.6	40	4	AAF81504	Aaf81504 Novel hum
c 154	13.8	47.6	40	4	AAF87604	Aaf87604 DNA assoc
c 155	13.8	47.6	44	9	ADD19669	Add19669 Oreochrom
c 156	13.8	47.6	50	3	AAA69485	Aaa69485 EMP-Fc fu
157	13.8	47.6	50	4	AAL28559	Aal28559 Human SNP
c 158	13.8	47.6	50	4	AAL29197	Aal29197 Human SNP
c 159	13.8	47.6	50	6	ABL35727	Abl35727 EPO mimet
c 160	13.8	47.6	51	4	AAL30984	Aal30984 Human SNP
c 161	13.8	47.6	51	4	AAL32426	Aal32426 Human SNP
c 162	13.8	47.6	51	4	AAL30985	Aal30985 Human SNP
c 163	13.8	47.6	51	4	AAI73350	Aai73350 Human sil
164	13.8	47.6	52	2	AAX79606	Aax79606 Capture e
165	13.8	47.6	55	7	ACC41834	Acc41834 Pre-contr
166	13.8	47.6	57	3	AAA69497	Aaa69497 EMP-EMP-F
167	13.8	47.6	57	6	ABL35739	Abl35739 EMP-EMP g
c 168	13.8	47.6	60	3	AAA69498	Aaa69498 EMP-EMP-F
c 169	13.8	47.6	60	6	ABL35740	Abl35740 EMP-EMP g
170	13.8	47.6	60	6	ABN38801	Abn38801 Human spl
171	13.8	47.6	61	3	AAA69484	Aaa69484 EMP-Fc fu
172	13.8	47.6	61	6	ABL35726	Abl35726 EPO mimet
c 173	13.8	47.6	65	6	ABN57018	Abn57018 Mouse spl
174	13.6	46.9	25	8	ACK28090	Ack28090 Human mic
175	13.6	46.9	25	8	ACK28091	Ack28091 Human mic
c 176	13.6	46.9	25	8	ACH52985	Ach52985 DNA targe
177	13.6	46.9	29	2	AAX59360	Aax59360 Human sec
178	13.6	46.9	29	4	AAS59328	Aas59328 Human sec
179	13.6	46.9	29	6	ABA90997	Aba90997 Biotinyla
c 180	13.6	46.9	31	7	ABZ62333	Abz62333 Human K-R
c 181	13.6	46.9	39	3	AAZ93063	Aaz93063 Primer fo
182	13.6	46.9	41	3	AAC81610	Aac81610 Enterobac
183	13.6	46.9	41	5	AAC88466	Aac88466 Helper ol
184	13.6	46.9	41	6	ABS60026	Abs60026 Human DNA
c 185	13.6	46.9	47	2	AAV64247	Aav64247 Plasmid p
c 186	13.6	46.9	47	3	AAZ96994	Aaz96994 S. cerevi
c 187	13.6	46.9	47	3	AAZ67496	Aaz67496 Human map
188	13.6	46.9	49	6	ABZ48800	Abz48800 Human ald
189	13.6	46.9	49	6	ABZ46289	Abz46289 Human ald
190	13.6	46.9	49	7	ACA54531	Aca54531 T7 termin
c 191	13.6	46.9	50	6	ABZ07629	Abz07629 Human leu
192	13.6	46.9	51	4	AAL28191	Aal28191 Human SNP
193	13.6	46.9	52	2	AAV69331	Aav69331 Human LIR
194	13.6	46.9	52	3	AAA54594	Aaa54594 Primer us
c 195	13.6	46.9	52	5	AAH81488	Aah81488 Plasmid p
c 196	13.6	46.9	52	7	ACA54530	Aca54530 T7 termin
c 197	13.6	46.9	52	7	ABX13507	Abx13507 Expressio
198	13.6	46.9	60	5	AAH81489	Aah81489 Plasmid p
199	13.6	46.9	60	6	ABN33662	Abn33662 Human spl
c 200	13.6	46.9	60	6	ABN38340	Abn38340 Human spl
201	13.6	46.9	60	7	ABX13508	Abx13508 Expressio
202	13.6	46.9	65	6	ABZ28242	Abz28242 Candida g
c 203	13.6	46.9	65	6	ABN55335	Abn55335 Mouse spl
c 204	13.6	46.9	65	6	ABN29388	Abn29388 Rat splic
c 205	13.6	46.9	65	6	ABN55942	Abn55942 Mouse spl
206	13.6	46.9	69	6	ABK88416	Abk88416 DNA encod
c 207	13.6	46.9	76	4	AAH36317	Aah36317 Human col
c 208	13.6	46.9	79	2	AAT57902	Aat57902 L-selecti
c 209	13.6	46.9	79	8	ADA21866	Ada21866 HGF trunc
210	13.4	46.2	25	3	AAC96310	Aac96310 HLA DPB1
211	13.4	46.2	25	3	AAC95712	Aac95712 HLA DQA1
212	13.4	46.2	25	3	AAC96748	Aac96748 HLA HLA-A
213	13.4	46.2	25	3	AAC95762	Aac95762 HLA DRB1
214	13.4	46.2	25	3	AAC95694	Aac95694 HLA DPB1
215	13.4	46.2	25	3	AAC96616	Aac96616 HLA DRB34
216	13.4	46.2	25	3	AAC96695	Aac96695 HLA HLA-A
217	13.4	46.2	25	3	AAC96425	Aac96425 HLA DQA1
218	13.4	46.2	25	3	AAC95707	Aac95707 HLA DQA1
219	13.4	46.2	25	3	AAC96406	Aac96406 HLA DQA1
220	13.4	46.2	25	3	AAC96284	Aac96284 HLA DPB1
221	13.4	46.2	25	3	AAC95693	Aac95693 HLA DPB1
222	13.4	46.2	25	3	AAC96404	Aac96404 HLA DQA1
223	13.4	46.2	25	3	AAC96739	Aac96739 HLA HLA-A
c 224	13.4	46.2	29	5	AAF16760	Aaf16760 Mazie ace
c 225	13.4	46.2	31	7	ABZ66333	Abz66333 Human HER
226	13.4	46.2	32	2	AAT38321	Aat38321 Human B-c
227	13.4	46.2	32	2	AAV32214	Aav32214 Human B-c
228	13.4	46.2	32	3	AAZ90109	Aaz90109 PCR prime
229	13.4	46.2	32	4	AAD09309	Aad09309 3' primer
230	13.4	46.2	32	6	AAD39315	Aad39315 Human BTG
231	13.4	46.2	34	6	ABN89825	Abn89825 Physcomit
232	13.4	46.2	34	6	ABN89846	Abn89846 Physcomit
c 233	13.4	46.2	34	7	ABS58345	Abs58345 HIV TAT g
c 234	13.4	46.2	40	3	AAA69500	Aaa69500 EMP-EMP-F
c 235	13.4	46.2	40	6	ABL35742	Abl35742 EMP-EMP g
236	13.4	46.2	41	6	ABZ45049	Abz45049 Human ATP
237	13.4	46.2	41	6	ABZ47644	Abz47644 Human ATP
c 238	13.4	46.2	41	7	ABZ57065	Abz57065 Excitator
c 239	13.4	46.2	43	6	ABZ27929	Abz27929 Candida e
c 240	13.4	46.2	46	2	AAQ69295	Aaq69295 Human atr
c 241	13.4	46.2	46	2	AAQ69292	Aaq69292 Human pro
c 242	13.4	46.2	46	2	AAT63757	Aat63757 Human atr

c 243	13.4	46.2	46	2	AAT63754	Aat63754 Human pro
c 244	13.4	46.2	46	2	AAX17045	Aax17045 Test sequ
c 245	13.4	46.2	46	2	AAX17042	Aax17042 Test sequ
c 246	13.4	46.2	46	6	ABK82536	Abk82536 DNA bindi
c 247	13.4	46.2	46	6	ABK82533	Abk82533 DNA bindi
c 248	13.4	46.2	46	10	ADE80075	Ade80075 Duplex ol
c 249	13.4	46.2	46	10	ADE80072	Ade80072 Duplex ol
250	13.4	46.2	47	3	AAZ68495	Aaz68495 Human map
c 251	13.4	46.2	47	6	ABK40908	Abk40908 Human obe
c 252	13.4	46.2	50	2	AAQ69296	Aaq69296 Human atr
253	13.4	46.2	50	2	AAT16219	Aat16219 Bryodin 1
c 254	13.4	46.2	50	2	AAT63758	Aat63758 Human atr
c 255	13.4	46.2	50	2	AAX17046	Aax17046 Test sequ
c 256	13.4	46.2	50	6	ABK82537	Abk82537 DNA bindi
c 257	13.4	46.2	50	6	AAS20796	Aas20796 Clostridi
c 258	13.4	46.2	50	6	AAS20794	Aas20794 Clostridi
c 259	13.4	46.2	50	6	AAS20795	Aas20795 Clostridi
c 260	13.4	46.2	50	10	ADE80076	Ade80076 Duplex ol
c 261	13.4	46.2	51	4	AAH89568	Aah89568 Human ATP
c 262	13.4	46.2	51	4	AAH89570	Aah89570 Human ATP
c 263	13.4	46.2	51	6	ABK93908	Abk93908 Human imm
c 264	13.4	46.2	60	4	AAI66246	Aai66246 S cerevis
c 265	13.4	46.2	60	6	ABN34937	Abn34937 Human spl
266	13.4	46.2	60	6	ABN46743	Abn46743 Human spl
c 267	13.4	46.2	60	6	ABN45568	Abn45568 Human spl
c 268	13.4	46.2	60	6	ABN34903	Abn34903 Human spl
c 269	13.4	46.2	60	9	ADD02782	Add02782 S. cerevi
270	13.4	46.2	61	3	AAC11143	Aac11143 Human sec
271	13.4	46.2	65	6	ABN29039	Abn29039 Rat splic
c 272	13.4	46.2	65	6	ABN56629	Abn56629 Mouse spl
c 273	13.4	46.2	65	6	ABN30536	Abn30536 Rat splic
274	13.4	46.2	65	6	ABN28913	Abn28913 Rat splic
c 275	13.4	46.2	79	3	AAZ34474	Aaz34474 Selenocys
276	13.2	45.5	22	4	AAI65515	Aai65515 PCR prime
277	13.2	45.5	22	6	AAS20552	Aas20552 Human uro
278	13.2	45.5	24	6	ABL53929	Abl53929 Human dih
c 279	13.2	45.5	25	2	AAV38001	Aav38001 SCEPO sec
c 280	13.2	45.5	25	6	ABS75571	Abs75571 Human PAP
c 281	13.2	45.5	25	6	ABS75579	Abs75579 Human PAP
282	13.2	45.5	25	8	ACI70679	Aci70679 Human mic
c 283	13.2	45.5	27	2	AAT35597	Aat35597 M.lufu pr
c 284	13.2	45.5	27	2	AAV24380	Aav24380 Target se
c 285	13.2	45.5	27	2	AAV12963	Aav12963 Mycobacte
c 286	13.2	45.5	27	9	ADD24591	Add24591 DNA polym
287	13.2	45.5	28	2	AAV65776	Aav65776 Helicobac
288	13.2	45.5	29	3	AAA04144	Aaa04144 Polymorph
c 289	13.2	45.5	29	3	AAA03920	Aaa03920 Polymorph
c 290	13.2	45.5	30	2	AAQ70868	Aaq70868 Target se
c 291	13.2	45.5	31	6	ABK21414	Abk21414 Human ERG
c 292	13.2	45.5	31	7	ABZ63565	Abz63565 Human H-R
c 293	13.2	45.5	31	7	ACD65682	Acd65682 HCV minus
c 294	13.2	45.5	32	2	AAT74999	Aat74999 DNA polym
295	13.2	45.5	37	6	ABK10624	Abk10624 Forward P
296	13.2	45.5	38	2	AAT53863	Aat53863 Rat ICAM
297	13.2	45.5	38	2	AAT81819	Aat81819 Human c-m
298	13.2	45.5	41	6	ABZ49819	Abz49819 Human car
299	13.2	45.5	41	6	ABZ45787	Abz45787 Human car
c 300	13.2	45.5	45	3	AAA58441	Aaa58441 Human Fac
301	13.2	45.5	47	3	AAZ68034	Aaz68034 Human map
c 302	13.2	45.5	50	4	AAL29812	Aal29812 Human SNP
303	13.2	45.5	50	4	AAL30642	Aal30642 Human SNP
304	13.2	45.5	50	6	ABZ07540	Abz07540 Human leu
305	13.2	45.5	50	6	ABZ07167	Abz07167 Human leu
c 306	13.2	45.5	50	6	ABZ06077	Abz06077 Human leu
307	13.2	45.5	50	6	ABZ02031	Abz02031 Human leu
c 308	13.2	45.5	50	6	ABZ06777	Abz06777 Human leu
c 309	13.2	45.5	50	6	ABZ07461	Abz07461 Human leu
310	13.2	45.5	51	2	AAX15975	Aax15975 PCR prime
311	13.2	45.5	51	4	AAL29895	Aal29895 Human SNP
c 312	13.2	45.5	51	4	AAH79827	Aah79827 Human DNA
c 313	13.2	45.5	59	2	AAX15990	Aax15990 PCR prime
314	13.2	45.5	60	6	ABN46195	Abn46195 Human spl
315	13.2	45.5	60	6	ABN33411	Abn33411 Human spl
316	13.2	45.5	60	6	ABN37230	Abn37230 Human spl
317	13.2	45.5	60	6	ABN36430	Abn36430 Human spl
c 318	13.2	45.5	60	6	ABN43565	Abn43565 Human spl
c 319	13.2	45.5	60	6	ABN34767	Abn34767 Human spl
c 320	13.2	45.5	64	2	AAT17547	Aat17547 T. litora
321	13.2	45.5	65	6	ABN53393	Abn53393 Mouse spl
c 322	13.2	45.5	65	6	ABN57539	Abn57539 Mouse spl
c 323	13.2	45.5	66	2	AAV77042	Aav77042 Staphyloc
324	13.2	45.5	66	2	AAV44966	Aav44966 Primer (-
c 325	13.2	45.5	69	9	ADC87772	Adc87772 F. culmor
326	13.2	45.5	73	3	AAZ99748	Aaz99748 Sense PCR
c 327	13	44.8	17	9	ADB40246	Adb40246 Tumour su
c 328	13	44.8	24	8	ACF04473	Acf04473 Real time
329	13	44.8	25	3	AAC96238	Aac96238 16s rRNA
330	13	44.8	25	3	AAC96061	Aac96061 16s rRNA
331	13	44.8	25	3	AAC96035	Aac96035 16s rRNA
332	13	44.8	25	3	AAC96472	Aac96472 HLA DQB1
c 333	13	44.8	27	2	AAV94045	Aav94045 Human IL-
c 334	13	44.8	27	2	AAV94525	Aav94525 Canine IL
335	13	44.8	27	3	AAZ57934	Aaz57934 Petunia n
336	13	44.8	27	3	AAZ50222	Aaz50222 Reverse p
337	13	44.8	30	2	AAZ25845	Aaz25845 Human pol
c 338	13	44.8	30	6	ABS63309	Abs63309 Ehrlichia
c 339	13	44.8	31	2	AAX38771	Aax38771 Human gen
340	13	44.8	31	3	AAA79150	Aaa79150 Human gen
c 341	13	44.8	31	4	ABK06279	Abk06279 Human NOG
c 342	13	44.8	36	4	AAC82688	Aac82688 P. chryso
c 343	13	44.8	38	4	ABK05323	Abk05323 Human NOG
344	13	44.8	47	3	AAZ65749	Aaz65749 Human map
345	13	44.8	49	2	AAT29867	Aat29867 Human pap
346	13	44.8	50	4	AAL30175	Aal30175 Human SNP
347	13	44.8	50	4	AAL31756	Aal31756 Human SNP
348	13	44.8	50	4	AAL32315	Aal32315 Human SNP
349	13	44.8	50	4	AAL30176	Aal30176 Human SNP
350	13	44.8	50	4	AAL32316	Aal32316 Human SNP
c 351	13	44.8	50	5	ABL00071	Abl00071 Human sil
c 352	13	44.8	50	5	ABL00072	Abl00072 Human sil
353	13	44.8	50	6	ABZ00884	Abz00884 Human leu
c 354	13	44.8	50	6	ABZ08097	Abz08097 Human leu
355	13	44.8	51	4	AAL28858	Aal28858 Human SNP
356	13	44.8	51	4	AAL28857	Aal28857 Human SNP
357	13	44.8	51	4	AAL31866	Aal31866 Human SNP
358	13	44.8	51	4	AAI76934	Aai76934 Human sil
c 359	13	44.8	54	2	AAQ48758	Aaq48758 TDH (144-
c 360	13	44.8	59	3	AAC17294	Aac17294 Human sec
361	13	44.8	60	6	ABN44575	Abn44575 Human spl
c 362	13	44.8	60	6	ABN40354	Abn40354 Human spl
c 363	13	44.8	60	6	ABN48250	Abn48250 Human spl
364	13	44.8	60	6	ABN48512	Abn48512 Human spl
c 365	13	44.8	60	6	ABN42516	Abn42516 Human spl
366	13	44.8	60	6	ABN33033	Abn33033 Human spl
c 367	13	44.8	60	6	ABN49443	Abn49443 Human spl
c 368	13	44.8	60	6	ABN41672	Abn41672 Human spl
c 369	13	44.8	64	8	ACC85362	Acc85362 N tabacum
c 370	13	44.8	64	9	ADE34205	Ade34205 Tobacco P
c 371	13	44.8	65	6	ABZ26554	Abz26554 Candida e
372	13	44.8	65	6	ABZ29362	Abz29362 Candida g
c 373	13	44.8	65	6	ABN31437	Abn31437 Rat splic
c 374	13	44.8	65	6	ABN51150	Abn51150 Mouse spl
c 375	13	44.8	65	6	ABN52191	Abn52191 Mouse spl
376	13	44.8	70	2	AAT78754	Aat78754 SELEX gen
377	13	44.8	73	3	AAC15012	Aac15012 Human sec
378	13	44.8	74	1	AAN92746	Aan92746 Tobamo ty
c 379	13	44.8	77	3	AAC11488	Aac11488 Human sec
380	13	44.8	78	6	ABL36079	Abl36079 M. jannas
381	13	44.8	78	6	ABL35948	Abl35948 M. jannas
c 382	12.8	44.1	16	4	AAS56770	Aas56770 BR2 prote
c 383	12.8	44.1	16	4	AAS56813	Aas56813 Target va
c 384	12.8	44.1	18	3	AAZ71110	Aaz71110 Human bia
c 385	12.8	44.1	20	3	AAZ38550	Aaz38550 Human mic
c 386	12.8	44.1	20	4	AAF76513	Aaf76513 Human EFE
c 387	12.8	44.1	20	7	ABZ92389	Abz92389 Human oli
c 388	12.8	44.1	21	2	AAZ26365	Aaz26365 Human pol

	389	12.8	44.1	22	6	ABX09317	Abx09317 Arteriosc
c	390	12.8	44.1	22	6	ABX09315	Abx09315 Arteriosc
c	391	12.8	44.1	23	2	AAZ33129	Aaz33129 Treponema
	392	12.8	44.1	24	6	ABX09316	Abx09316 Arteriosc
	393	12.8	44.1	24	6	ABS61529	Abs61529 Analyte s
c	394	12.8	44.1	24	7	ACA88954	Aca88954 Selection
	395	12.8	44.1	25	3	AAC96573	Aac96573 HLA DRB34
	396	12.8	44.1	29	3	AAF03788	Aaf03788 Hammerhea
	397	12.8	44.1	30	2	AAT39014	Aat39014 Interleuk
c	398	12.8	44.1	30	2	AAV32703	Aav32703 Human chr
c	399	12.8	44.1	31	4	ABK08758	Abk08758 Human CD2
c	400	12.8	44.1	31	4	ABK08843	Abk08843 Human CD2
c	401	12.8	44.1	31	7	ACD43699	Acd43699 Human gen
c	402	12.8	44.1	31	7	ABZ65562	Abz65562 Human HER
c	403	12.8	44.1	31	7	ACD64894	Acd64894 HCV minus
c	404	12.8	44.1	31	7	ACD58247	Acd58247 HCV DNAzy
	405	12.8	44.1	32	3	AAC58314	Aac58314 Human PRO
c	406	12.8	44.1	33	2	AAQ84754	Aaq84754 Primer (P
c	407	12.8	44.1	33	2	AAV62864	Aav62864 Primer fo
c	408	12.8	44.1	33	2	AAV81677	Aav81677 Oligonucl
	409	12.8	44.1	33	8	AAL51909	Aal51909 Human typ
	410	12.8	44.1	33	8	ABZ79531	Abz79531 Human typ
	411	12.8	44.1	36	2	AAT54510	Aat54510 Human IL-
c	412	12.8	44.1	38	8	AAL57621	Aal57621 VL revers
c	413	12.8	44.1	38	9	ADD35948	Add35948 Single ch
	414	12.8	44.1	40	2	AAT69456	Aat69456 Plasmid p
	415	12.8	44.1	40	2	AAX88880	Aax88880 Circular
	416	12.8	44.1	40	6	ABA98155	Aba98155 Oligonucl
	417	12.8	44.1	40	6	ABQ96125	Abq96125 Tumour su
c	418	12.8	44.1	41	2	AAV51170	Aav51170 Maize pol
	419	12.8	44.1	41	6	ABZ70107	Abz70107 Human rib
c	420	12.8	44.1	41	6	ABZ49197	Abz49197 Human ald
c	421	12.8	44.1	41	6	ABZ49196	Abz49196 Human ald
c	422	12.8	44.1	43	3	AAZ36463	Aaz36463 PCR prime
	423	12.8	44.1	48	6	ABZ47549	Abz47549 Human ATP
	424	12.8	44.1	50	4	AAL29317	Aal29317 Human SNP
	425	12.8	44.1	50	4	AAL27907	Aal27907 Human SNP
	426	12.8	44.1	50	4	AAL28929	Aal28929 Human SNP
	427	12.8	44.1	50	6	ABZ04870	Abz04870 Human leu
c	428	12.8	44.1	50	6	ABZ01104	Abz01104 Human leu
c	429	12.8	44.1	50	6	ABZ01829	Abz01829 Human leu
c	430	12.8	44.1	50	6	ABZ02897	Abz02897 Human leu
c	431	12.8	44.1	50	6	ABZ05954	Abz05954 Human leu
	432	12.8	44.1	50	9	ADD93363	Add93363 Flt1 gene
	433	12.8	44.1	51	4	AAL29743	Aal29743 Human SNP
c	434	12.8	44.1	51	4	AAL31074	Aal31074 Human SNP
	435	12.8	44.1	51	4	AAL29742	Aal29742 Human SNP
	436	12.8	44.1	51	4	AAL30586	Aal30586 Human SNP
	437	12.8	44.1	51	4	AAL28931	Aal28931 Human SNP
c	438	12.8	44.1	51	4	AAL33262	Aal33262 Human SNP
	439	12.8	44.1	53	2	AAV55867	Aav55867 Plasmid v
	440	12.8	44.1	53	6	ABA98834	Aba98834 SSU oligo
	441	12.8	44.1	54	3	AAZ38935	Aaz38935 hCAT1 bin
	442	12.8	44.1	54	3	AAZ38741	Aaz38741 HCAT1 bin
c	443	12.8	44.1	60	6	ABN47809	Abn47809 Human spl
	444	12.8	44.1	60	6	ABN33663	Abn33663 Human spl
c	445	12.8	44.1	60	6	ABN38057	Abn38057 Human spl
	446	12.8	44.1	60	6	ABN40220	Abn40220 Human spl
	447	12.8	44.1	60	6	ABN36863	Abn36863 Human spl
c	448	12.8	44.1	60	6	ABN59534	Abn59534 Human spl
c	449	12.8	44.1	60	7	AAD47406	Aad47406 Oligo 1F5
c	450	12.8	44.1	62	6	ABL36038	Abl36038 D. melano
c	451	12.8	44.1	63	2	AAQ26745	Aaq26745 VL-PAL fu
	452	12.8	44.1	65	6	AAK98759	Aak98759 Yeast flo
	453	12.8	44.1	65	6	ABZ27314	Abz27314 Candida e
c	454	12.8	44.1	65	6	ABZ27078	Abz27078 Candida e
	455	12.8	44.1	65	6	ABN28506	Abn28506 Rat splic
c	456	12.8	44.1	65	6	ABN27823	Abn27823 Rat splic
	457	12.8	44.1	65	6	ABN54631	Abn54631 Mouse spl
	458	12.8	44.1	65	6	ABN58232	Abn58232 Mouse spl
c	459	12.8	44.1	65	6	ABN28655	Abn28655 Rat splic
	460	12.8	44.1	65	6	ABN56610	Abn56610 Mouse spl
	461	12.8	44.1	69	2	AAV30083	Aav30083 Probe use
c	462	12.8	44.1	69	3	AAC12591	Aac12591 Human sec
	463	12.8	44.1	69	3	AAC16727	Aac16727 Human sec
	464	12.8	44.1	69	4	AAD21265	Aad21265 Interleuk
	465	12.8	44.1	69	7	ABX96959	Abx96959 Interleuk
	466	12.8	44.1	69	8	ABX80098	Abx80098 Cytokine
	467	12.8	44.1	71	2	AAQ79043	Aaq79043 SREBP-1 p
	468	12.8	44.1	71	4	AAL19895	Aal19895 Human bre
	469	12.8	44.1	72	2	AAT19449	Aat19449 Human gen
c	470	12.8	44.1	73	2	AAT21249	Aat21249 Human gen
c	471	12.8	44.1	77	4	AAH99422	Aah99422 Human pro
	472	12.8	44.1	79	3	AAC15512	Aac15512 Human sec
	473	12.6	43.4	19	6	ABK41500	Abk41500 Human CTN
c	474	12.6	43.4	20	4	AAF44608	Aaf44608 Novel mou
	475	12.6	43.4	21	4	AAF97622	Aaf97622 Human gen
c	476	12.6	43.4	23	3	AAA27822	Aaa27822 North Ame
	477	12.6	43.4	25	2	AAV26343	Aav26343 Human pro
	478	12.6	43.4	25	2	AAX26090	Aax26090 Prostate
	479	12.6	43.4	25	3	AAZ87576	Aaz87576 Primer sp
	480	12.6	43.4	25	4	AAS03993	Aas03993 Biomarker
	481	12.6	43.4	25	8	ACI05194	Aci05194 Human mic
	482	12.6	43.4	25	8	ACK12687	Ack12687 Human mic
	483	12.6	43.4	25	8	ACI76452	Aci76452 Human mic
	484	12.6	43.4	26	2	AAX29472	Aax29472 A. tumefa
	485	12.6	43.4	27	6	ABL99419	Abl99419 Right PCR
c	486	12.6	43.4	29	3	AAA04581	Aaa04581 Polymorph
c	487	12.6	43.4	29	3	AAA15377	Aaa15377 PCR prime
	488	12.6	43.4	30	2	AAQ30958	Aaq30958 GAD probe
	489	12.6	43.4	30	2	AAX24222	Aax24222 Oligonucl
	490	12.6	43.4	30	3	AAA88006	Aaa88006 Mycobacte
c	491	12.6	43.4	31	4	ABL48014	Abl48014 Human GRI
c	492	12.6	43.4	31	6	ABK21284	Abk21284 Human ERG
c	493	12.6	43.4	31	6	ABK21784	Abk21784 Human ERG
c	494	12.6	43.4	31	7	ACA08357	Aca08357 Necrosis
c	495	12.6	43.4	31	7	ABZ66155	Abz66155 Human HER
c	496	12.6	43.4	31	7	ABZ66420	Abz66420 Human HER
c	497	12.6	43.4	31	7	ACD65563	Acd65563 HCV minus
c	498	12.6	43.4	33	5	AAI68374	Aai68374 Human lig
c	499	12.6	43.4	34	2	AAV28476	Aav28476 GM-CSF re
c	500	12.6	43.4	34	2	AAX91824	Aax91824 Porphorym
c	501	12.6	43.4	36	2	AAQ90693	Aaq90693 B. burgdo
	502	12.6	43.4	36	2	AAT69812	Aat69812 Heparin c
c	503	12.6	43.4	36	5	AAH45219	Aah45219 Human zin
c	504	12.6	43.4	36	6	ABK47813	Abk47813 Borrelia
c	505	12.6	43.4	37	2	AAX91988	Aax91988 Porphorym
c	506	12.6	43.4	38	8	AAL57619	Aal57619 VL revers
c	507	12.6	43.4	38	8	AAL57612	Aal57612 VL revers
c	508	12.6	43.4	38	9	ADD35939	Add35939 Single ch
c	509	12.6	43.4	38	9	ADD35946	Add35946 Single ch
c	510	12.6	43.4	41	6	ABZ47620	Abz47620 Human ATP
	511	12.6	43.4	42	2	AAX27115	Aax27115 Primer P4
	512	12.6	43.4	42	2	AAX27116	Aax27116 Primer P5
	513	12.6	43.4	42	2	AAX84825	Aax84825 PCR prime
c	514	12.6	43.4	42	3	AAA71123	Aaa71123 Molecular
c	515	12.6	43.4	42	3	AAA71131	Aaa71131 Molecular
c	516	12.6	43.4	45	2	AAQ69592	Aaq69592 Human gen
c	517	12.6	43.4	45	2	AAQ69380	Aaq69380 Human fib
c	518	12.6	43.4	45	2	AAT64054	Aat64054 Human fib
c	519	12.6	43.4	45	2	AAT63842	Aat63842 Human fib
c	520	12.6	43.4	45	2	AAX17130	Aax17130 Test sequ
c	521	12.6	43.4	45	2	AAX17342	Aax17342 Test sequ
c	522	12.6	43.4	45	6	ABK82833	Abk82833 DNA bindi
c	523	12.6	43.4	45	6	ABK82621	Abk82621 DNA bindi
c	524	12.6	43.4	45	10	ADE80160	Ade80160 Duplex ol
c	525	12.6	43.4	45	10	ADE80372	Ade80372 Duplex ol
	526	12.6	43.4	46	3	AAA71097	Aaa71097 Molecular
c	527	12.6	43.4	47	3	AAZ68609	Aaz68609 Human map
c	528	12.6	43.4	47	7	ACC45113	Acc45113 T7 termin
c	529	12.6	43.4	48	2	AAX86907	Aax86907 T7 termin
c	530	12.6	43.4	49	2	AAX55485	Aax55485 Oligonucl
	531	12.6	43.4	49	4	AAD15365	Aad15365 Oligo 1F,
c	532	12.6	43.4	49	4	AAD15366	Aad15366 Oligo 1A,
c	533	12.6	43.4	49	6	AAD27055	Aad27055 T7 transc
c	534	12.6	43.4	50	2	AAQ53287	Aaq53287 T7 gene 1

c	535	12.6	43.4	50	2	AAQ69593	Aaq69593	Human gen		608	12.4	42.8	22	2	AAV99612	Aav99612	Maize clp
c	536	12.6	43.4	50	2	AAT64055	Aat64055	Human fib	c	609	12.4	42.8	22	4	AAH02036	Aah02036	S. pneumo
c	537	12.6	43.4	50	2	AAX17343	Aax17343	Test sequ		610	12.4	42.8	23	2	AAV44902	Aav44902	Petunia f
	538	12.6	43.4	50	3	AAZ29462	Aaz29462	PCR Prime	c	611	12.4	42.8	23	9	ADE14269	Ade14269	Optineuri
	539	12.6	43.4	50	4	AAL28877	Aal28877	Human SNP		612	12.4	42.8	24	2	AAQ13644	Aaq13644	Intron/ex
c	540	12.6	43.4	50	4	AAI77621	Aai77621	Human sil	c	613	12.4	42.8	25	2	AAV90566	Aav90566	Forward P
c	541	12.6	43.4	50	4	AAH23403	Aah23403	PCR metho		614	12.4	42.8	25	3	AAC96107	Aac96107	16s rRNA
	542	12.6	43.4	50	4	AAF76755	Aaf76755	T flavus		615	12.4	42.8	25	8	ACI60360	Aci60360	Human mic
c	543	12.6	43.4	50	6	ABK82834	Abk82834	DNA bindi	c	616	12.4	42.8	25	8	ACI33391	Aci33391	Human mic
	544	12.6	43.4	50	6	ABZ00862	Abz00862	Human leu		617	12.4	42.8	25	8	ACI43897	Aci43897	Human mic
c	545	12.6	43.4	50	6	ABZ04583	Abz04583	Human leu		618	12.4	42.8	25	8	ACK21623	Ack21623	Human mic
c	546	12.6	43.4	50	6	ABZ08111	Abz08111	Human leu		619	12.4	42.8	25	8	ACI04888	Aci04888	Human mic
	547	12.6	43.4	50	6	ABZ06512	Abz06512	Human leu	c	620	12.4	42.8	25	8	ACK13887	Ack13887	Human mic
c	548	12.6	43.4	50	6	ABZ06902	Abz06902	Human leu		621	12.4	42.8	27	1	AAN92527	Aan92527	Tag seque
	549	12.6	43.4	50	9	ADD69694	Add69694	Nucleic a		622	12.4	42.8	28	2	AAV26787	Aav26787	BRCA1 gen
c	550	12.6	43.4	50	10	ADE80373	Ade80373	Duplex ol		623	12.4	42.8	28	6	ABQ78708	Abq78708	PCR prime
	551	12.6	43.4	51	2	AAV70234	Aav70234	scFv anti	c	624	12.4	42.8	30	2	AAT14104	Aat14104	Reverse p
	552	12.6	43.4	51	4	AAL27621	Aal27621	Human SNP	c	625	12.4	42.8	30	2	AAT73980	Aat73980	Rat THY-1
	553	12.6	43.4	51	4	AAL28669	Aal28669	Human SNP	c	626	12.4	42.8	30	2	AAV23325	Aav23325	Rat THY-1
	554	12.6	43.4	51	4	AAH89375	Aah89375	Human pro	c	627	12.4	42.8	30	2	AAX34786	Aax34786	Protein P
	555	12.6	43.4	54	3	AAA73943	Aaa73943	GFP Ile(A	c	628	12.4	42.8	30	7	AAL50155	Aal50155	Truncated
	556	12.6	43.4	55	3	AAZ22764	Aaz22764	Oligonucl		629	12.4	42.8	31	2	AAV42111	Aav42111	Petunia f
c	557	12.6	43.4	56	1	AAN70018	Aan70018	Sequence		630	12.4	42.8	31	2	AAX84616	Aax84616	PCR prime
c	558	12.6	43.4	56	6	ABZ06465	Abz06465	Human leu	c	631	12.4	42.8	31	4	ABK06469	Abk06469	Human NOG
c	559	12.6	43.4	56	6	ABZ07524	Abz07524	Human leu		632	12.4	42.8	32	2	AAV44058	Aav44058	NEO-r gen
	560	12.6	43.4	56	6	ABZ06761	Abz06761	Human leu		633	12.4	42.8	32	2	AAZ07122	Aaz07122	SV40 prom
	561	12.6	43.4	56	6	ABZ07445	Abz07445	Human leu	c	634	12.4	42.8	32	6	AAS19194	Aas19194	DRalpha/m
c	562	12.6	43.4	56	6	ABZ07151	Abz07151	Human leu		635	12.4	42.8	33	2	AAZ07732	Aaz07732	C. boidin
	563	12.6	43.4	56	6	ABZ06855	Abz06855	Human leu		636	12.4	42.8	34	2	AAX78755	Aax78755	Human RAN
c	564	12.6	43.4	59	3	AAC29059	Aac29059	Human sec		637	12.4	42.8	34	2	AAX78758	Aax78758	Human RAN
c	565	12.6	43.4	60	2	AAQ68079	Aaq68079	Lysozyme		638	12.4	42.8	37	2	AAQ78589	Aaq78589	Vector am
	566	12.6	43.4	60	6	ABN38458	Abn38458	Human spl		639	12.4	42.8	37	2	AAT80878	Aat80878	Vector am
	567	12.6	43.4	60	6	ABN35929	Abn35929	Human spl		640	12.4	42.8	37	2	AAX26012	Aax26012	Internal
c	568	12.6	43.4	60	6	ABN47870	Abn47870	Human spl		641	12.4	42.8	37	3	AAI64536	Aai64536	EB virus
c	569	12.6	43.4	60	6	ABN47051	Abn47051	Human spl	c	642	12.4	42.8	38	4	ABK08054	Abk08054	Human CD2
c	570	12.6	43.4	60	6	ABN58644	Abn58644	Human spl	c	643	12.4	42.8	39	9	ADE25934	Ade25934	GalNAc-tr
c	571	12.6	43.4	60	6	ABN40128	Abn40128	Human spl		644	12.4	42.8	40	4	AAF90359	Aaf90359	A probe f
c	572	12.6	43.4	60	6	ABN38979	Abn38979	Human spl	c	645	12.4	42.8	41	2	AAV50867	Aav50867	Maize pol
	573	12.6	43.4	62	2	AAQ63648	Aaq63648	Antisense	c	646	12.4	42.8	41	2	AAV47725	Aav47725	Maize pol
	574	12.6	43.4	62	2	AAT45374	Aat45374	Antisense		647	12.4	42.8	41	6	ABK11369	Abk11369	NADH dehy
	575	12.6	43.4	62	2	AAX05886	Aax05886	Ribozyme	c	648	12.4	42.8	41	6	ABZ47613	Abz47613	Human ATP
c	576	12.6	43.4	62	6	AAD27057	Aad27057	Oligo oGN		649	12.4	42.8	41	7	AAL55402	Aal55402	Ribosomal
c	577	12.6	43.4	63	6	ABV83599	Abv83599	Human bre		650	12.4	42.8	42	2	AAV29706	Aav29706	PCR prime
	578	12.6	43.4	65	6	AAD27059	Aad27059	Oligo oGN	c	651	12.4	42.8	44	2	AAQ78588	Aaq78588	Vector am
c	579	12.6	43.4	65	6	AAD27058	Aad27058	Oligo oGN	c	652	12.4	42.8	44	2	AAT80877	Aat80877	Vector am
c	580	12.6	43.4	65	6	ABZ28932	Abz28932	Candida g		653	12.4	42.8	44	3	AAC64475	Aac64475	Plasmid p
	581	12.6	43.4	65	6	ABN27408	Abn27408	Rat splic		654	12.4	42.8	44	9	ADC17368	Adc17368	3' primer
	582	12.6	43.4	65	6	ABN31044	Abn31044	Rat splic		655	12.4	42.8	45	4	AAH43433	Aah43433	Primer J
c	583	12.6	43.4	65	6	ABN53625	Abn53625	Mouse spl		656	12.4	42.8	46	7	ACC59633	Acc59633	SV40 earl
	584	12.6	43.4	65	6	ABN52800	Abn52800	Mouse spl		657	12.4	42.8	46	7	ACC59670	Acc59670	SV40 earl
c	585	12.6	43.4	66	2	AAX55972	Aax55972	Nucleotid		658	12.4	42.8	46	8	ACC59772	Acc59772	SV40 earl
c	586	12.6	43.4	67	3	AAC70120	Aac70120	TGF-beta-		659	12.4	42.8	46	8	ACC85111	Acc85111	SV40 earl
c	587	12.6	43.4	69	2	AAV64780	Aav64780	HIV anti-		660	12.4	42.8	47	3	AAZ69403	Aaz69403	Human map
c	588	12.6	43.4	69	4	AAH24792	Aah24792	Nucleotid		661	12.4	42.8	48	4	AAF57995	Aaf57995	SV40 prom
	589	12.6	43.4	69	7	ACF42757	Acf42757	shMOG gen	c	662	12.4	42.8	49	6	ABQ96309	Abq96309	Tumour su
c	590	12.6	43.4	69	7	ACA74098	Aca74098	Hepatitis	c	663	12.4	42.8	50	4	AAL31422	Aal31422	Human SNP
	591	12.6	43.4	70	6	AAD27056	Aad27056	Oligo oGN	c	664	12.4	42.8	50	4	AAL31421	Aal31421	Human SNP
c	592	12.6	43.4	71	3	AAC69900	Aac69900	TGF-beta-	c	665	12.4	42.8	50	4	AAL31423	Aal31423	Human SNP
c	593	12.6	43.4	71	3	AAC69899	Aac69899	TGF-beta-	c	666	12.4	42.8	50	4	AAL31428	Aal31428	Human SNP
	594	12.6	43.4	72	2	AAQ63646	Aaq63646	Sense pri		667	12.4	42.8	50	4	AAI78469	Aai78469	Human sil
	595	12.6	43.4	72	2	AAT45372	Aat45372	Sense pri		668	12.4	42.8	50	6	ABZ04580	Abz04580	Human leu
	596	12.6	43.4	72	2	AAX05884	Aax05884	Ribozyme	c	669	12.4	42.8	50	6	ABZ00701	Abz00701	Human leu
c	597	12.6	43.4	72	6	ABL35979	Abl35979	C. elegan	c	670	12.4	42.8	50	6	ABZ05299	Abz05299	Human leu
	598	12.6	43.4	74	7	ABZ70378	Abz70378	Streptavi	c	671	12.4	42.8	50	6	ABZ07777	Abz07777	Human leu
	599	12.6	43.4	74	7	AAL56603	Aal56603	Oligonucl		672	12.4	42.8	51	2	AAZ31115	Aaz31115	RNA ligan
c	600	12.6	43.4	76	3	AAC18253	Aac18253	Human sec		673	12.4	42.8	51	3	AAA77185	Aaa77185	Human clo
	601	12.6	43.4	76	6	AAD25663	Aad25663	Oligonucl	c	674	12.4	42.8	51	4	AAL31424	Aal31424	Human SNP
	602	12.6	43.4	76	6	AAD25561	Aad25561	HCV genom	c	675	12.4	42.8	51	4	AAL27454	Aal27454	Human SNP
c	603	12.6	43.4	77	2	AAT06139	Aat06139	HIV-1 int		676	12.4	42.8	51	4	AAL28846	Aal28846	Human SNP
c	604	12.6	43.4	78	3	AAC58720	Aac58720	Taq polym		677	12.4	42.8	51	4	AAL28845	Aal28845	Human SNP
c	605	12.6	43.4	80	9	ADD69704	Add69704	Nucleic a	c	678	12.4	42.8	51	4	AAL31427	Aal31427	Human SNP
c	606	12.4	42.8	17	7	ACC51245	Acc51245	Human tum	c	679	12.4	42.8	51	4	AAL26911	Aal26911	Human SNP
c	607	12.4	42.8	17	9	ADB44357	Adb44357	Tumour su		680	12.4	42.8	51	4	AAI78468	Aai78468	Human sil

	681	12.4	42.8	51	4	AAI74954 Aai74954 Human sil
c	682	12.4	42.8	51	4	AAI77812 Aai77812 Human sil
c	683	12.4	42.8	51	4	AAH38324 Aah38324 Human SNP
c	684	12.4	42.8	53	2	AAV55866 Aav55866 Plasmid v
	685	12.4	42.8	53	6	ABQ94596 Abq94596 Tumour su
c	686	12.4	42.8	53	6	ABA98833 Aba98833 SSU oligo
	687	12.4	42.8	54	6	ABI82118 Abi82118 BRCA2 mut
c	688	12.4	42.8	54	7	ABX93131 Abx93131 Human pho
	689	12.4	42.8	60	6	ABQ78142 Abq78142 Synthetic
c	690	12.4	42.8	60	6	ABN37841 Abn37841 Human spl
c	691	12.4	42.8	60	6	ABN48412 Abn48412 Human spl
c	692	12.4	42.8	60	6	ABN32595 Abn32595 Human spl
c	693	12.4	42.8	60	8	ACC83938 Acc83938 GPBP-inte
c	694	12.4	42.8	64	6	ABV89209 Abv89209 Human col
	695	12.4	42.8	65	4	AAF25853 Aaf25853 Fusion pr
c	696	12.4	42.8	65	6	ABZ29669 Abz29669 Candida g
c	697	12.4	42.8	65	6	ABZ26279 Abz26279 Candida e
c	698	12.4	42.8	65	6	ABN52800 Abn52800 Mouse spl
c	699	12.4	42.8	66	1	AAN97097 Aan97097 Sequence
c	700	12.4	42.8	66	1	AAN97093 Aan97093 Sequence
c	701	12.4	42.8	66	1	AAN97077 Aan97077 Sequence
c	702	12.4	42.8	66	2	AAX18036 Aax18036 C. tracho
c	703	12.4	42.8	66	2	AAX18040 Aax18040 C. tracho
c	704	12.4	42.8	66	6	ABK97753 Abk97753 C. tracho
c	705	12.4	42.8	66	6	ABK97769 Abk97769 C. tracho
c	706	12.4	42.8	66	6	ABK97773 Abk97773 C. tracho
	707	12.4	42.8	68	2	AAV77073 Aav77073 Staphyloc
c	708	12.4	42.8	70	9	ADC64809 Adc64809 HuBBK-4H-
c	709	12.4	42.8	71	1	AAN92747 Aan92747 Tobacco m
	710	12.4	42.8	71	6	AAL49309 Aal49309 Glut tran
	711	12.4	42.8	73	6	ABA94609 Aba94609 Nucleotid
	712	12.4	42.8	78	2	AAX11553 Aax11553 Human bia
c	713	12.2	42.1	17	6	ABS74757 Abs74757 Human PAP
c	714	12.2	42.1	17	6	ABS74755 Abs74755 Human PAP
c	715	12.2	42.1	17	6	ABS74756 Abs74756 Human PAP
c	716	12.2	42.1	17	6	ABK56306 Abk56306 Human CLC
c	717	12.2	42.1	19	2	AAT63463 Aat63463 Antisense
	718	12.2	42.1	19	4	AAI64794 Aai64794 Human MAP
c	719	12.2	42.1	20	2	AAV34493 Aav34493 BRCA1 exo
	720	12.2	42.1	20	2	AAZ00488 Aaz00488 Human thi
c	721	12.2	42.1	20	2	AAX97056 Aax97056 PCR prime
c	722	12.2	42.1	20	7	ABZ90675 Abz90675 Human oli
	723	12.2	42.1	20	9	ADE15806 Ade15806 Human PKR
c	724	12.2	42.1	21	6	ABL45287 Abl45287 Human chr
c	725	12.2	42.1	21	6	ABS97814 Abs97814 Human NAD
c	726	12.2	42.1	21	6	ABS97813 Abs97813 Human NAD
c	727	12.2	42.1	21	9	ADE77850 Ade77850 DNA oligo
	728	12.2	42.1	22	6	ABX97518 Abx97518 Human NOV
	729	12.2	42.1	22	9	ADE15289 Ade15289 Transcrip
	730	12.2	42.1	23	6	ABA04436 Aba04436 Human SP1
c	731	12.2	42.1	24	6	ABA04165 Aba04165 Human int
c	732	12.2	42.1	24	6	ABL56597 Abl56597 PCR prime
	733	12.2	42.1	24	6	AAS18594 Aas18594 Human ald
	734	12.2	42.1	24	6	ABS61836 Abs61836 Analyte s
	735	12.2	42.1	25	3	AAC96576 Aac96576 HLA DRB34
	736	12.2	42.1	25	3	AAC96177 Aac96177 16s rRNA
c	737	12.2	42.1	25	3	AAZ37265 Aaz37265 PCR prime
c	738	12.2	42.1	25	3	AAZ49605 Aaz49605 PCR prime
	739	12.2	42.1	25	3	AAZ49606 Aaz49606 PCR prime
	740	12.2	42.1	25	4	AAH40259 Aah40259 SNP speci
c	741	12.2	42.1	25	6	ABS75570 Abs75570 Human PAP
c	742	12.2	42.1	25	6	ABS75580 Abs75580 Human PAP
	743	12.2	42.1	25	7	ACF03537 Acf03537 Brassica
c	744	12.2	42.1	25	8	ACI83822 Aci83822 Human mic
	745	12.2	42.1	25	8	ACI65128 Aci65128 Human mic
	746	12.2	42.1	25	8	ACI17964 Aci17964 Human mic
c	747	12.2	42.1	25	8	ACI61329 Aci61329 Human mic
	748	12.2	42.1	25	8	ACI18911 Aci18911 Human mic
	749	12.2	42.1	26	6	ABK89530 Abk89530 Barley ye
	750	12.2	42.1	27	3	AAK99530 Aak99530 Alpha-hum
	751	12.2	42.1	27	6	ABK52619 Abk52619 Minority
c	752	12.2	42.1	27	9	ADE77852 Ade77852 DNA oligo
	753	12.2	42.1	28	6	AAK98658 Aak98658 Regulator
	754	12.2	42.1	28	6	AAL44152 Aal44152 Barley ye
c	755	12.2	42.1	29	2	AAT51441 Aat51441 Mouse hea
c	756	12.2	42.1	29	2	AAQ83497 Aaq83497 MAb 3B9 g
c	757	12.2	42.1	29	2	AAT34103 Aat34103 Mouse gam
c	758	12.2	42.1	29	2	AAV03498 Aav03498 Mouse gam
c	759	12.2	42.1	29	2	AAX85897 Aax85897 Heavy cha
c	760	12.2	42.1	29	2	AAX79522 Aax79522 PCR prime
c	761	12.2	42.1	29	3	AAA04123 Aaa04123 Polymorph
	762	12.2	42.1	30	2	AAQ51374 Aaq51374 Chlamydia
	763	12.2	42.1	30	2	AAQ51383 Aaq51383 Chlamydia
	764	12.2	42.1	30	2	AAT41651 Aat41651 Endo-xylo
	765	12.2	42.1	30	2	AAT99549 Aat99549 Cucumber
	766	12.2	42.1	30	2	AAZ07507 Aaz07507 Cucumber
c	767	12.2	42.1	30	4	AAD14197 Aad14197 Mouse int
c	768	12.2	42.1	30	5	AAF82178 Aaf82178 Human ADA
	769	12.2	42.1	30	6	ABK14828 Abk14828 Rat FGF-R
c	770	12.2	42.1	30	6	ABX69591 Abx69591 Novel Hel
	771	12.2	42.1	30	7	ABX14304 Abx14304 PCR prime
c	772	12.2	42.1	30	9	ADC15694 Adc15694 E. intest
c	773	12.2	42.1	30	9	ADC89712 Adc89712 L. johnso
c	774	12.2	42.1	31	4	ABK08907 Abk08907 Human CD2
c	775	12.2	42.1	31	7	ABZ66489 Abz66489 Human HER
c	776	12.2	42.1	31	7	ABZ65632 Abz65632 Human HER
c	777	12.2	42.1	31	7	ABZ65797 Abz65797 Human HER
c	778	12.2	42.1	31	7	ABZ63733 Abz63733 Human H-R
c	779	12.2	42.1	31	7	ACD64924 Acd64924 HCV minus
c	780	12.2	42.1	31	7	ACD65686 Acd65686 HCV minus
c	781	12.2	42.1	31	7	ACD65462 Acd65462 HCV minus
	782	12.2	42.1	32	3	AAZ90002 Aaz90002 PCR prime
	783	12.2	42.1	32	5	AAS44007 Aas44007 Neisseria
c	784	12.2	42.1	32	6	ABA90805 Aba90805 Bacillus
	785	12.2	42.1	32	6	ABS69861 Abs69861 Ehrlichia
	786	12.2	42.1	32	6	AAS98184 Aas98184 Human gra
c	787	12.2	42.1	32	8	ACD27683 Acd27683 M. hyopne
	788	12.2	42.1	33	2	AAQ49317 Aaq49317 Degenerin
c	789	12.2	42.1	33	6	ABL54520 Abl54520 Pectinatu
	790	12.2	42.1	34	2	AAQ10833 Aaq10833 Pneumocys
	791	12.2	42.1	34	2	AAT42445 Aat42445 Probe 149
	792	12.2	42.1	34	3	AAC60374 Aac60374 Primer CA
	793	12.2	42.1	34	3	AAD61870 Aad61870 Human imm
	794	12.2	42.1	34	7	ABX14302 Abx14302 Human his
	795	12.2	42.1	36	2	AAV16746 Aav16746 Human Wnt
	796	12.2	42.1	36	5	AAF24068 Aaf24068 Wnt gene
	797	12.2	42.1	37	9	ADD15476 Add15476 PCR prime
	798	12.2	42.1	38	2	AAT81739 Aat81739 Human c-m
	799	12.2	42.1	38	2	AAT82034 Aat82034 Human c-m
	800	12.2	42.1	38	2	AAT82070 Aat82070 Human c-m
	801	12.2	42.1	38	7	ACD53775 Acd53775 HBV zinzy
c	802	12.2	42.1	38	8	AAL57617 Aal57617 VL revers
c	803	12.2	42.1	38	9	ADD35944 Add35944 Single ch
	804	12.2	42.1	39	2	AAT61103 Aat61103 Chimeric
c	805	12.2	42.1	39	2	AAT61104 Aat61104 Chimeric
c	806	12.2	42.1	39	2	AAV90775 Aav90775 Primer Y1
c	807	12.2	42.1	39	5	AAF82147 Aaf82147 T7 termin
	808	12.2	42.1	40	3	AAZ32756 Aaz32756 Chimeric
c	809	12.2	42.1	40	6	ABL54532 Abl54532 Pectinatu
c	810	12.2	42.1	41	2	AAV61543 Aav61543 Primer, u
c	811	12.2	42.1	41	6	ABS60715 Abs60715 Human DNA
c	812	12.2	42.1	41	6	ABZ49930 Abz49930 Human NQO
	813	12.2	42.1	41	6	ABZ43999 Abz43999 Human glu
c	814	12.2	42.1	41	6	ABZ43459 Abz43459 Human NQO
	815	12.2	42.1	41	6	ABZ49439 Abz49439 Human glu
	816	12.2	42.1	41	6	ABZ45003 Abz45003 Human ald
	817	12.2	42.1	41	6	ABZ49264 Abz49264 Human ald
	818	12.2	42.1	41	7	ABQ77331 Abq77331 Bovine H-
	819	12.2	42.1	41	7	AAL50786 Aal50786 Exonuclea
c	820	12.2	42.1	41	7	ABZ57438 Abz57438 Amidase f
c	821	12.2	42.1	41	8	AAD56833 Aad56833 HIV type
c	822	12.2	42.1	42	3	AAA57825 Aaa57825 Oligonucl
	823	12.2	42.1	42	7	ABQ77326 Abq77326 Bovine H-
	824	12.2	42.1	42	7	AAL50781 Aal50781 Exonuclea
	825	12.2	42.1	42	8	AAD56832 Aad56832 HIV type
c	826	12.2	42.1	43	6	ABZ26815 Abz26815 Candida e

827	12.2	42.1	43	7	ABT17588	Abt17588	Invader d
828	12.2	42.1	43	9	ADD15475	Add15475	PCR prime
c 829	12.2	42.1	44	2	AAT58442	Aat58442	Beta fibr
c 830	12.2	42.1	44	2	AAV82552	Aav82552	Probe FIB
831	12.2	42.1	45	2	AAQ55414	Aaq55414	Antifunga
c 832	12.2	42.1	45	2	AAT42486	Aat42486	3' primer
c 833	12.2	42.1	45	3	AAA65020	Aaa65020	Oligonucl
834	12.2	42.1	45	3	AAA65034	Aaa65034	Oligonucl
c 835	12.2	42.1	47	3	AAZ67939	Aaz67939	Human map
836	12.2	42.1	47	3	AAZ68139	Aaz68139	Human map
837	12.2	42.1	48	4	AAF29316	Aaf29316	Primer ba
c 838	12.2	42.1	48	6	ABZ47051	Abz47051	Human ATP
c 839	12.2	42.1	48	7	ABZ25347	Abz25347	PCR prime
840	12.2	42.1	48	8	AAD57383	Aad57383	Human 2H9
841	12.2	42.1	49	2	AAQ23172	Aaq23172	HPV probe
c 842	12.2	42.1	49	8	ADA94805	Ada94805	Primer P0
c 843	12.2	42.1	50	4	AAL31223	Aal31223	Human SNP
844	12.2	42.1	50	4	AAI76933	Aai76933	Human sil
845	12.2	42.1	50	4	AAI77143	Aai77143	Human sil
c 846	12.2	42.1	50	4	AAI77355	Aai77355	Human sil
847	12.2	42.1	50	4	AAI76773	Aai76773	Human sil
c 848	12.2	42.1	50	6	ABZ44775	Abz44775	Human ATP
849	12.2	42.1	50	6	ABZ01370	Abz01370	Human leu
c 850	12.2	42.1	50	6	ABZ08007	Abz08007	Human leu
851	12.2	42.1	50	6	ABZ00408	Abz00408	Human leu
c 852	12.2	42.1	50	6	ABZ01788	Abz01788	Human leu
c 853	12.2	42.1	50	6	ABZ02646	Abz02646	Human leu
c 854	12.2	42.1	50	6	ABZ01645	Abz01645	Human leu
855	12.2	42.1	50	6	ABZ00057	Abz00057	Human leu
856	12.2	42.1	50	6	ABZ03582	Abz03582	Human leu
c 857	12.2	42.1	50	6	ABZ04133	Abz04133	Human leu
c 858	12.2	42.1	50	6	ABZ04512	Abz04512	Human leu
859	12.2	42.1	51	4	AAL27139	Aal27139	Human SNP
c 860	12.2	42.1	51	4	AAL30469	Aal30469	Human SNP
c 861	12.2	42.1	51	4	AAL30723	Aal30723	Human SNP
c 862	12.2	42.1	51	4	AAL27471	Aal27471	Human SNP
c 863	12.2	42.1	51	4	AAL33251	Aal33251	Human SNP
c 864	12.2	42.1	51	4	AAI77354	Aai77354	Human sil
c 865	12.2	42.1	51	4	AAI73351	Aai73351	Human sil
866	12.2	42.1	51	4	AAI76851	Aai76851	Human sil
867	12.2	42.1	51	4	AAI73954	Aai73954	Human sil
868	12.2	42.1	51	4	AAI76772	Aai76772	Human sil
869	12.2	42.1	51	4	AAI76932	Aai76932	Human sil
c 870	12.2	42.1	51	4	AAI73636	Aai73636	Human sil
871	12.2	42.1	51	4	AAH37844	Aah37844	Human SNP
872	12.2	42.1	51	4	AAH40664	Aah40664	Human SNP
c 873	12.2	42.1	51	4	AAH40260	Aah40260	Human SNP
874	12.2	42.1	51	4	AAH79850	Aah79850	Human DNA
c 875	12.2	42.1	51	6	ABI98852	Abi98852	Oligonucl
876	12.2	42.1	53	7	ABV77344	Abv77344	SSA4-forw
877	12.2	42.1	54	3	AAZ38931	Aaz38931	hCAT1 bin
878	12.2	42.1	54	3	AAZ38737	Aaz38737	HCAT1 bin
879	12.2	42.1	54	3	AAC24508	Aac24508	Human sec
c 880	12.2	42.1	54	5	AAF83721	Aaf83721	CMV-DNA t
881	12.2	42.1	54	6	ABZ50607	Abz50607	Human car
882	12.2	42.1	55	2	AAQ33929	Aaq33929	Sequence
883	12.2	42.1	55	3	AAA57826	Aaa57826	Oligonucl
c 884	12.2	42.1	55	7	ACF19156	Acf19156	Tumour ce
c 885	12.2	42.1	55	9	ADC84979	Adc84979	MCF-7 bre
c 886	12.2	42.1	57	6	AAS97118	Aas97118	Humanised
c 887	12.2	42.1	57	7	AAL60533	Aal60533	MOD1R63 P
888	12.2	42.1	58	3	AAC94687	Aac94687	Cat flea
c 889	12.2	42.1	59	2	AAQ47050	Aaq47050	GM-CSF ol
c 890	12.2	42.1	59	2	AAQ67686	Aaq67686	Native B.
c 891	12.2	42.1	59	2	AAV30041	Aav30041	Oligonucl
892	12.2	42.1	59	3	AAA11243	Aaa11243	Primer P2
c 893	12.2	42.1	59	4	AAD21222	Aad21222	Immunomod
c 894	12.2	42.1	59	7	ABX96917	Abx96917	Immunomod
c 895	12.2	42.1	59	8	ABX80056	Abx80056	Human imm
896	12.2	42.1	60	2	AAT20453	Aat20453	Human gen
c 897	12.2	42.1	60	6	ABN42457	Abn42457	Human spl
c 898	12.2	42.1	60	6	ABN47117	Abn47117	Human spl
c 899	12.2	42.1	60	6	ABN42592	Abn42592	Human spl
c 900	12.2	42.1	60	6	ABN34401	Abn34401	Human spl
c 901	12.2	42.1	60	6	ABN46627	Abn46627	Human spl
902	12.2	42.1	60	6	ABN41771	Abn41771	Human spl
c 903	12.2	42.1	60	6	ABN43420	Abn43420	Human spl
c 904	12.2	42.1	60	6	ABN32208	Abn32208	Human spl
c 905	12.2	42.1	60	6	ABN42368	Abn42368	Human spl
c 906	12.2	42.1	60	6	ABN33329	Abn33329	Human spl
907	12.2	42.1	60	6	ABN40487	Abn40487	Human spl
c 908	12.2	42.1	60	6	ABN35456	Abn35456	Human spl
c 909	12.2	42.1	60	6	ABN49510	Abn49510	Human spl
c 910	12.2	42.1	60	6	ABN45393	Abn45393	Human spl
911	12.2	42.1	60	6	ABN49155	Abn49155	Human spl
c 912	12.2	42.1	60	9	ACF57704	Acf57704	DNA encod
c 913	12.2	42.1	60	9	ACF57703	Acf57703	DNA encod
914	12.2	42.1	62	3	AAZ96886	Aaz96886	S. cerevi
c 915	12.2	42.1	62	9	ADD80885	Add80885	DNA media
916	12.2	42.1	62	9	ADD80888	Add80888	DNA media
917	12.2	42.1	63	7	ABQ77328	Abq77328	Bovine H-
918	12.2	42.1	63	7	AAL50783	Aal50783	Exonuclea
919	12.2	42.1	64	2	AAX37724	Aax37724	Human PRO
920	12.2	42.1	65	6	ABZ29826	Abz29826	Candida g
921	12.2	42.1	65	6	ABZ29688	Abz29688	Candida g
922	12.2	42.1	65	6	ABZ28703	Abz28703	Candida g
923	12.2	42.1	65	6	ABZ27232	Abz27232	Candida e
c 924	12.2	42.1	65	6	ABN54908	Abn54908	Mouse spl
c 925	12.2	42.1	65	6	ABN53130	Abn53130	Mouse spl
c 926	12.2	42.1	65	6	ABN56298	Abn56298	Mouse spl
927	12.2	42.1	65	6	ABN55121	Abn55121	Mouse spl
c 928	12.2	42.1	65	6	ABN56758	Abn56758	Mouse spl
c 929	12.2	42.1	66	4	AAH36754	Aah36754	Human col
c 930	12.2	42.1	66	7	ABZ37238	Abz37238	Human lam
931	12.2	42.1	69	3	AAC32441	Aac32441	Human sec
932	12.2	42.1	70	2	AAT78724	Aat78724	SELEX gen
933	12.2	42.1	72	7	ABQ77332	Abq77332	Bovine H-
934	12.2	42.1	72	7	AAL50787	Aal50787	Exonuclea
c 935	12.2	42.1	72	8	ADA73647	Ada73647	Carcinoma
c 936	12.2	42.1	72	8	ADA02101	Ada02101	Mouse Fgf
c 937	12.2	42.1	72	9	ADB71840	Adb71840	Mouse car
c 938	12.2	42.1	73	7	ABT17591	Abt17591	Invader d
c 939	12.2	42.1	73	7	ABT17592	Abt17592	Invader d
c 940	12.2	42.1	74	5	AAF98688	Aaf98688	Human ova
941	12.2	42.1	75	2	AAT25666	Aat25666	Human gen
c 942	12.2	42.1	75	6	ABA82796	Aba82796	Human pro
c 943	12.2	42.1	78	2	AAQ47925	Aaq47925	cDNA frag
c 944	12.2	42.1	78	2	AAT50934	Aat50934	Mouse p53
c 945	12.2	42.1	80	2	AAT42515	Aat42515	Sequence
946	12	41.4	15	6	ABK67889	Abk67889	Human ADH
c 947	12	41.4	18	6	AAL49052	Aal49052	Drosophil
c 948	12	41.4	19	3	AAA83806	Aaa83806	cdk-we-hu
c 949	12	41.4	19	5	AAH58968	Aah58968	Cdk-we-hu
c 950	12	41.4	20	2	AAZ06182	Aaz06182	PCR prime
951	12	41.4	20	3	AAA40957	Aaa40957	Human TNF
c 952	12	41.4	20	7	ABZ98522	Abz98522	Human ICA
c 953	12	41.4	20	7	ABZ92390	Abz92390	Human oli
954	12	41.4	20	8	ACD05185	Acd05185	Tumour ne
955	12	41.4	21	3	AAZ72421	Aaz72421	Human bia
c 956	12	41.4	21	4	AAF23944	Aaf23944	Bacillus
957	12	41.4	21	6	ABQ79874	Abq79874	Nucleotid
c 958	12	41.4	21	6	ABX97345	Abx97345	Human NOV
959	12	41.4	22	2	AAX78215	Aax78215	E. rhusio
960	12	41.4	22	9	ADE15296	Ade15296	Transcrip
961	12	41.4	22	9	ADE15298	Ade15298	Transcrip
962	12	41.4	23	2	AAQ80875	Aaq80875	Dengue 1
963	12	41.4	23	6	ABS63220	Abs63220	Identific
c 964	12	41.4	23	9	ADD43577	Add43577	Oligonucl
c 965	12	41.4	24	6	ABX13855	Abx13855	Human rib
966	12	41.4	25	3	AAC95781	Aac95781	HLA DRB1
967	12	41.4	25	3	AAC96831	Aac96831	HLA HLA-C
968	12	41.4	25	3	AAC96194	Aac96194	16s rRNA
969	12	41.4	25	3	AAC96279	Aac96279	HLA DPB1
970	12	41.4	25	4	AAH38767	Aah38767	SNP speci
971	12	41.4	25	8	ACI47170	Aci47170	Human mic
972	12	41.4	25	8	ACI15377	Aci15377	Human mic

c 973	12	41.4	25	8	ACI45090	Aci45090 Human mic
974	12	41.4	25	8	ACI01313	Aci01313 Human mic
c 975	12	41.4	25	8	ACI20237	Aci20237 Human mic
c 976	12	41.4	25	8	ACK10194	Ack10194 Human mic
c 977	12	41.4	25	8	ACI80480	Aci80480 Human mic
c 978	12	41.4	25	8	ACI01290	Aci01290 Human mic
979	12	41.4	25	8	ACI05489	Aci05489 Human mic
c 980	12	41.4	25	8	ACI30316	Aci30316 Human mic
c 981	12	41.4	25	8	ACI20236	Aci20236 Human mic
c 982	12	41.4	25	8	ACK29363	Ack29363 Human mic
c 983	12	41.4	25	8	ACH52859	Ach52859 DNA targe
c 984	12	41.4	30	1	AAN80286	Aan80286 Sequence
c 985	12	41.4	30	1	AAN80276	Aan80276 Sequence
c 986	12	41.4	30	4	AAH91564	Aah91564 Human inf
987	12	41.4	30	9	ADD89904	Add89904 PCR prime
c 988	12	41.4	31	4	AAH97251	Aah97251 Human Chk
c 989	12	41.4	31	4	ABK06396	Abk06396 Human NOG
c 990	12	41.4	31	6	ABK21758	Abk21758 Human ERG
c 991	12	41.4	31	6	ABK21517	Abk21517 Human ERG
c 992	12	41.4	31	6	ABK21720	Abk21720 Human ERG
c 993	12	41.4	31	7	ACA08386	Aca08386 Necrosis
c 994	12	41.4	31	7	ACA08379	Aca08379 Necrosis
995	12	41.4	31	7	ACD54142	Acd54142 HBV DNAzy
c 996	12	41.4	31	7	ACD65121	Acd65121 HCV minus
c 997	12	41.4	31	7	ACD58897	Acd58897 HCV DNAzy
c 998	12	41.4	31	7	ACD60376	Acd60376 HCV DNAzy
c 999	12	41.4	31	7	ACD57571	Acd57571 HCV DNAzy
c1000	12	41.4	31	7	ACD60035	Acd60035 HCV DNAzy

ALIGNMENTS

RESULT 1
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29

RESULT 2
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds

CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

```
  Query Match             100.0%;  Score 29;  DB 3;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||||||||||||||||||||||||||
Db          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
```

RESULT 3
AAA71121
ID AAA71121 standard; DNA; 42 BP.
XX
AC AAA71121;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #127.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

```
  Query Match             100.0%;  Score 29;  DB 3;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.0028;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||::|:::::|:|||||||||||||:
Db          4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32
```

RESULT 4
AAA71128
ID AAA71128 standard; RNA; 42 BP.
XX
AC AAA71128;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #197.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA

CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 5
AAA71120
ID AAA71120 standard; DNA; 42 BP.
XX
AC AAA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 69.0%; Pred. No. 0.0028;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||::|:::::|:|||||||||||||:
Db 4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32

RESULT 6
AAA71116
ID AAA71116 standard; RNA; 42 BP.
XX
AC AAA71116;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #192.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to

CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 7
AAA71115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 8
AAA71129
ID AAA71129 standard; RNA; 42 BP.
XX
AC AAA71129;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #198.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24

CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 9
AAA70826
ID AAA70826 standard; RNA; 45 BP.
XX
AC AAA70826;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #26.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 222; Page 232; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 45 BP; 14 A; 7 C; 9 G; 0 T; 15 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 18 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 45

RESULT 10
AAA70825
ID AAA70825 standard; RNA; 45 BP.
XX
AC AAA70825;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #25.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 221; Page 232; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first

CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 45 BP; 14 A; 7 C; 9 G; 0 T; 15 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 18 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 45

RESULT 11
AAA71089
ID AAA71089 standard; DNA; 46 BP.
XX
AC AAA71089;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #112.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 71.4%; Pred. No. 0.0081;
Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
|||||::|:::::|:||||||||||||
Db 19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46

RESULT 12
AAA71106
ID AAA71106 standard; RNA; 46 BP.
XX
AC AAA71106;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #182.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second

CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46

RESULT 13
AAA71107
ID AAA71107 standard; RNA; 46 BP.
XX
AC AAA71107;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #183.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46

RESULT 14
AAA71088
ID AAA71088 standard; DNA; 46 BP.
XX
AC AAA71088;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #111.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4

CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 71.4%; Pred. No. 0.0081;
Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||::|:::::|:|||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46
```

RESULT 15
AAA71105
ID AAA71105 standard; RNA; 46 BP.
XX
AC AAA71105;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #181.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||||||||||||||||||||||||||
Db         19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46
```

RESULT 16
AAA71090
ID AAA71090 standard; DNA; 46 BP.
XX
AC AAA71090;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #113.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified

CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 71.4%; Pred. No. 0.0081;
Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||::|:::::|:|||||||||||||
Db 19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46

RESULT 17
AAA71113
ID AAA71113 standard; RNA; 42 BP.
XX
AC AAA71113;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #189.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 85.5%; Score 24.8; DB 3; Length 42;
Best Local Similarity 92.9%; Pred. No. 0.22;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
|||||||||||||||||||| | |||||
Db 5 AAGAUUCUUUUUGUAAGCCCUACGGGCU 32

RESULT 18
AAA71118
ID AAA71118 standard; DNA; 42 BP.
XX
AC AAA71118;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #124.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be

CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 U; 0 Other;

Query Match 85.5%; Score 24.8; DB 3; Length 42;
Best Local Similarity 60.7%; Pred. No. 0.22;
Matches 17; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||::|::::|:|||||| | ||||:
Db 5 AAGATTCTTTTTGTAAGCCCTACGGGCT 32

RESULT 19
AAA71126
ID AAA71126 standard; RNA; 42 BP.
XX
AC AAA71126;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #195.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 85.5%; Score 24.8; DB 3; Length 42;
Best Local Similarity 92.9%; Pred. No. 0.22;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||| | |||||
Db 5 AAGAUUCUUUUUGUAAGCCCUACGGGCU 32

RESULT 20
AAA71085
ID AAA71085 standard; DNA; 46 BP.
XX
AC AAA71085;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #108.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural

```
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 12 A; 7 C; 9 G; 18 T; 0 U; 0 Other;

  Query Match             82.1%;  Score 23.8;  DB 3;  Length 46;
  Best Local Similarity   63.0%;  Pred. No. 0.62;
  Matches   17;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||::|:::::|:|||||| | ||||
Db         20 AAGATTCTTTTTGTAAGCCCTACGGGC 46


RESULT 21
AAA71103
ID   AAA71103 standard; RNA; 46 BP.
XX
AC   AAA71103;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #179.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 12 A; 7 C; 9 G; 0 T; 18 U; 0 Other;

  Query Match             82.1%;  Score 23.8;  DB 3;  Length 46;
  Best Local Similarity   92.6%;  Pred. No. 0.62;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||||||||||||||||| | ||||
Db         20 AAGAUUCUUUUUGUAAGCCCUACGGGC 46


RESULT 22
AAA70828
ID   AAA70828 standard; RNA; 29 BP.
XX
AC   AAA70828;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #28.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Claim 235; Page 235; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
```

```
SQ   Sequence 29 BP; 5 A; 5 C; 7 G; 0 T; 12 U; 0 Other;

  Query Match             80.0%;  Score 23.2;  DB 3;  Length 29;
  Best Local Similarity   89.3%;  Pred. No. 1.1;
  Matches   25;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||||||||||||||| | |||||
Db          2 AUGAUUCUUUUUGUAAGCCCUAGGGGCU 29


RESULT 23
AAA71123
ID   AAA71123 standard; DNA; 42 BP.
XX
AC   AAA71123;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #129.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 U; 0 Other;

  Query Match             80.0%;  Score 23.2;  DB 3;  Length 42;
  Best Local Similarity   57.1%;  Pred. No. 1.1;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              | ||::|:::::|:|||||| | ||||:
Db          5 ATGATTCTTTTTGTAAGCCCTAGGGGCT 32


RESULT 24
AAA71131
ID   AAA71131 standard; RNA; 42 BP.
XX
AC   AAA71131;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #200.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 126; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 9 A; 6 C; 9 G; 0 T; 18 U; 0 Other;
```

```
  Query Match           80.0%;  Score 23.2;  DB 3;  Length 42;
  Best Local Similarity  89.3%;  Pred. No. 1.1;
  Matches   25;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              | ||||||||||||||||||| | |||||
Db          5 AUGAUUCUUUUUGUAAGCCCUAGGGGCU 32


RESULT 25
AAA70824
ID   AAA70824 standard; RNA; 45 BP.
XX
AC   AAA70824;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #24.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Claim 220; Page 232; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 45 BP; 11 A; 6 C; 9 G; 0 T; 19 U; 0 Other;

  Query Match           76.6%;  Score 22.2;  DB 3;  Length 45;
  Best Local Similarity  88.9%;  Pred. No. 3.2;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||||||||||||||||||| | ||||
Db         19 AUGAUUCUUUUUGUAAGCCCUAGGGGC 45


RESULT 26
AAA71087
ID   AAA71087 standard; DNA; 46 BP.
XX
AC   AAA71087;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #110.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

  Query Match           76.6%;  Score 22.2;  DB 3;  Length 46;
  Best Local Similarity  59.3%;  Pred. No. 3.2;
```

Matches 16; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||::|:::::|:|||||| | ||||
Db         20 ATGATTCTTTTTGTAAGCCCTAGGGGC 46
```

RESULT 27
AAA71096
ID AAA71096 standard; DNA; 46 BP.
XX
AC AAA71096;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #119.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 76.6%; Score 22.2; DB 3; Length 46;
Best Local Similarity 59.3%; Pred. No. 3.2;
Matches 16; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||::|:::::|:|||||| | ||||
Db         20 ATGATTCTTTTTGTAAGCCCTAGGGGC 46
```

RESULT 28
AAA71099
ID AAA71099 standard; DNA; 46 BP.
XX
AC AAA71099;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #122.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 76.6%; Score 22.2; DB 3; Length 46;
Best Local Similarity 59.3%; Pred. No. 3.2;
Matches 16; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||::|:::::|:|||||| | ||||
Db         20 ATGATTCTTTTTGTAAGCCCTAGGGGC 46
```

RESULT 29
AAA71100
ID AAA71100 standard; DNA; 46 BP.
XX
AC AAA71100;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #123.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

```
  Query Match            76.6%;  Score 22.2;  DB 3;  Length 46;
  Best Local Similarity  59.3%;  Pred. No. 3.2;
  Matches   16;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||::|:::::|:|||||| | ||||
Db         20 ATGATTCTTTTTGTAAGCCCTAGGGGC 46
```

RESULT 30
AAA71104
ID AAA71104 standard; RNA; 46 BP.
XX
AC AAA71104;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #180.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 0 T; 19 U; 0 Other;

```
  Query Match            76.6%;  Score 22.2;  DB 3;  Length 46;
  Best Local Similarity  88.9%;  Pred. No. 3.2;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | ||||||||||||||||||| | ||||
```

```
Db          20 AUGAUUCUUUUUGUAAGCCCUAGGGGC 46
```

```
RESULT 31
AAA71114
ID   AAA71114 standard; RNA; 42 BP.
XX
AC   AAA71114;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #190.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 8 C; 7 G; 0 T; 16 U; 0 Other;

  Query Match             73.1%;  Score 21.2;  DB 3;  Length 42;
  Best Local Similarity   88.5%;  Pred. No. 9;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGG 27
              |||||||||||||||||||| | | |
Db          5 AAGAUUCUUUUUGUAAGCCCUAGGCG 30
```

```
RESULT 32
AAA71119
ID   AAA71119 standard; DNA; 42 BP.
XX
AC   AAA71119;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #125.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 8 C; 7 G; 16 T; 0 U; 0 Other;

  Query Match             73.1%;  Score 21.2;  DB 3;  Length 42;
  Best Local Similarity   57.7%;  Pred. No. 9;
  Matches   15;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGG 27
              ||||::|:::::|:|||||| | | |
Db          5 AAGATTCTTTTTGTAAGCCCTAGGCG 30
```

```
RESULT 33
AAA71127
ID   AAA71127 standard; RNA; 42 BP.
XX
AC   AAA71127;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #196.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 126; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 8 C; 7 G; 0 T; 16 U; 0 Other;

  Query Match             73.1%;  Score 21.2;  DB 3;  Length 42;
  Best Local Similarity   88.5%;  Pred. No. 9;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGG 27
              |||||||||||||||||||| | | |
Db          5 AAGAUUCUUUUUGUAAGCCCUAGGCG 30
```

```
RESULT 34
AAA71094
ID   AAA71094 standard; DNA; 46 BP.
XX
AC   AAA71094;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #117.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 12 A; 7 C; 9 G; 18 T; 0 U; 0 Other;

  Query Match             73.1%;  Score 21.2;  DB 3;  Length 46;
  Best Local Similarity   57.7%;  Pred. No. 9.1;
  Matches   15;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGG 27
              ||||::|:::::|:|||||| | | |
Db         20 AAGATTCTTTTTGTAAGCCCTAGGCG 45
```

RESULT 35

```
AAA71110
ID   AAA71110 standard; RNA; 46 BP.
XX
AC   AAA71110;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #186.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 12 A; 7 C; 9 G; 0 T; 18 U; 0 Other;

  Query Match             73.1%;  Score 21.2;  DB 3;  Length 46;
  Best Local Similarity   88.5%;  Pred. No. 9.1;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGG 27
              |||||||||||||||||||| | | |
Db         20 AAGAUUCUUUUUGUAAGCCCUAGGCG 45


RESULT 36
AAA71098
ID   AAA71098 standard; DNA; 46 BP.
XX
AC   AAA71098;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #121.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 11 A; 5 C; 6 G; 17 T; 0 U; 7 Other;

  Query Match             69.0%;  Score 20;  DB 3;  Length 46;
  Best Local Similarity   60.0%;  Pred. No. 31;
  Matches   12;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCC 21
              ||||::|:::::|:||||||
Db         20 AAGATTCTTTTTGTAAGCCC 39


RESULT 37
AAA71102
ID   AAA71102 standard; RNA; 46 BP.
```

```
XX
AC   AAA71102;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #178.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 11 A; 5 C; 6 G; 0 T; 17 U; 7 Other;

  Query Match             69.0%;  Score 20;  DB 3;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCC 21
              ||||||||||||||||||||
Db         20 AAGAUUCUUUUUGUAAGCCC 39


RESULT 38
AAA71084
ID   AAA71084 standard; DNA; 46 BP.
XX
AC   AAA71084;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #107.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 11 A; 5 C; 6 G; 17 T; 0 U; 7 Other;

  Query Match             69.0%;  Score 20;  DB 3;  Length 46;
  Best Local Similarity   60.0%;  Pred. No. 31;
  Matches   12;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCC 21
              ||||::|:::::|:||||||
Db         20 AAGATTCTTTTTGTAAGCCC 39


RESULT 39
AAA71124
ID   AAA71124 standard; DNA; 42 BP.
XX
AC   AAA71124;
```

XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #130.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 11 A; 10 C; 7 G; 14 T; 0 U; 0 Other;

Query Match 67.6%; Score 19.6; DB 3; Length 42;
Best Local Similarity 57.7%; Pred. No. 47;
Matches 15; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```
Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||: |::: :|:|||||| | ||||:
Db          7 GATCCTTTCTGTAAGCCCTACGGGCT 32
```

RESULT 40
AAA71132
ID AAA71132 standard; RNA; 42 BP.
XX
AC AAA71132;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #201.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 11 A; 10 C; 7 G; 0 T; 14 U; 0 Other;

Query Match 67.6%; Score 19.6; DB 3; Length 42;
Best Local Similarity 84.6%; Pred. No. 47;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||| |||| |||||||||| | |||||
Db          7 GAUCCUUUCUGUAAGCCCUACGGGCU 32
```

Search completed: April 18, 2004, 08:04:18
Job time : 198.667 secs